VOL. XXII DECEMBER, 1938 No. 6

THE JOURNAL OF APPLIED PSYCHOLOGY

JAMES P. PORTER, *Editor*
OHIO UNIVERSITY ATHENS, OHIO

With a Board of Cooperating Editors
and
PAUL S. ACHILLES HENRY C. LINK
Associate Editors Representing
THE PSYCHOLOGICAL CORPORATION

SOME INTERESTING CONTRIBUTIONS IN THIS ISSUE:

CLARENCE LEUBA. Has Recent Research Undermined the Evidence for Extra Sensory Perception?

RICHARD S. SCHULTZ AND SYDNEY ROSLOW. Restandardization of the A-S Reaction Study as a Personnel Form

CLAUDE L. NEMZEK. The Value of the Bernreuter Personality Inventory for Direct and Differential Prediction of Academic Success as Measured by Teachers' Marks

NEAL DROUGHT. An Analysis of Eight Measures of Personality and Adjustment in Relation to Relative Scholastic Achievement

PUBLISHED BI-MONTHLY BY THE EDITOR
OHIO UNIVERSITY, ATHENS, OHIO

PRINTED BY THE SCIENCE PRESS PRINTING COMPANY, LANCASTER, PA.

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry May 7, 1935, at the post office at Lancaster, Pa. Copyright 1938 by James P. Porter.

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; M. E. Bolton, State Teachers College, San Diego Calif.; H. A. Bruce, Ohio State University; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. Peter W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Edward M. McConnell, Johns Hopkins University; Walter D. Scott, Northwestern University; G. R. Steadman, University of Iowa; Daniel Starch, New York; Lewis M. Terman, Leland Stanford Jr. University; Ir. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Denver, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business efficiency. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, comments, notices, books and offprints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by the Journals, in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $3.00 a year. Single numbers $1.25 to $2.00, depending upon the number of pages. Remittances and business communications should be addressed to the Publisher, James P. Porter, Ohio University, Athens, Ohio.

The Journal of
Applied Psychology

Edited by
JAMES P. PORTER
Ohio University
With a Board of Cooperating Editors

VOLUME XXII
1938

PUBLISHED BI-MONTHLY BY THE EDITOR
Ohio University, Athens, Ohio
Printed by
THE SCIENCE PRESS PRINTING COMPANY
N. Queen St. and McGovern Ave., Lancaster, Pa.

CONTENTS

THEODORA M. ABEL AND ROY HAMLIN. An Analysis of Capacities for Performance of Mental Defectives Skilled in Lace-making: II. Performance Pattern in Selected Non-Standardized Tasks 175

THEODORA M, ABEL: See Hamlin and Abel.

ROSE G. ANDERSON. A Note on a Case of Spelling Difficulty 211

ROSE G. ANDERSON. Some Technological Aspects of Counseling Adult Women 455

GEORGE K. BENNETT. A Simplified Scoring Method for the Bernreuter Personality Inventory 390

S. W. BIJOU. The Performance of Normal Children on the Randall's Island Performance Series 186

MARION A. BILLS. Relation of Scores in Strong's Interest Analysis Blanks to Success in Selling Casualty Insurance 97

ALBERT B. BLANKENSHIP AND HOWARD R. TAYLOR. Prediction of Vocational Proficiency in Three Machine Operations 518

RALPH R. BROWN AND VICTOR H. VOGEL. Psycho-physiological Reactions Following Painful Stimuli under Hypnotic Analgesia Contrasted with Gas Anaesthesia and Novocain Block 408

ALICE I. BRYAN AND RUTH E. PERL. A Comparison of Women Students Preparing for Three Different Vocations 161

HAROLD E. BURTT: See Mitchell and Burtt.

HERMAN G. CANADY. Sex Differences in Intelligence Among Negro College Freshmen 437

RALPH CHANNELL: See De Silva and Channell.

HERMAN A. COPELAND. Studies in the Reliability of Personnel Records 247

HORACE H. CORBIN, JR.: See Jenkins and Corbin.

VIRGINIA DAHLEM: See Greene and Dahlem.

CHARLES M. DAVIDSON. Permanency of Mental Alertness Measures 307

HARRY R. DE SILVA AND RALPH CHANNELL. Driver Clinics in the Field 59

ARTHUR F. DODGE. Social Dominance and Sales Personality 132

ARTHUR F. DODGE. What Are the Personality Traits of the Successful Sales-Person 229

NEAL DROUGHT. An Analysis of Eight Measures of Personality and Adjustment in Relation to Relative Scholastic Achievement 597

GEORGE J. DUDYCHA. The Estimation of Performance-Time in Simple Tasks 79

PAUL S. DWYER. An Analysis of Nineteen Occupational Scores on the Strong Vocational Interest Blank 8

ORAN W. EAGLESON. A Racial Comparison of Personality Traits 271

PHILIP EISENBERG. Judging Expressive Movement: I. Judgments of Sex and Dominance-Feeling from Handwriting Samples of Dominant and Non-Dominant Men and Women 480
PHILIP EISENBERG AND ELEANOR ZALOWITZ. Judging Expressive Movement: III. Judgments of Dominance-Feeling from Phonograph Records of Voice 620
ROYAL B. EMBREE, JR. The Use of Group Intelligence Tests with Correctional School Subjects 504
T. L. ENGLE. The Use of a Short Personality Schedule in High School Personnel Work 534
C. B. FISS: See Williamson, Stover and Fiss.
PHILIP FOWLER: See Wells, Williams and Fowler.
H. E. GEIGER, H. H. REMMERS AND R. J. GREENLY. Apprentice Attitudes Toward Their Training and the Construction of a Diagnostic Scale 32
CHARLES C. GIBBONS. A Short Scoring Method for the Link P. Q. Test 653
EDWARD B. GREENE AND VIRGINIA DAHLEM. The Michigan Vocational Preference List 558
R. J. GREENLY: See Geiger, Remmers and Greenly.
J. P. GUILFORD. A New Revision of the Army Alpha Examination and a Weighted Scoring for Three Primary Factors 239
ROY HAMLIN AND THEODORA M. ABEL. Test Pattern of Mental Defectives Skilled in Weaving 385
ROY HAMLIN: See Abel and Hamlin.
SAMUEL P. HAYES, JR. A Note on Personality and Family Position 347
JOHN G. JENKINS. Dependability of Psychological Brand Barometers: I. The Problem of Reliability 1
JOHN G. JENKINS AND HORACE H. CORBIN, JR. Dependability of Psychological Brand Barometers: II. The Problem of Validity 252
HAROLD A. KOHN: See Reymert and Kohn.
ARTHUR KOLSTAD. Employee Attitudes in a Department Store 470
JOSEPH S. KOPAS. The Point-Tally: A Modified Method of Scoring the Strong Vocational Interest Test 426
BARBARA KRUGER. A Statistical Analysis of the Humm-Wadsworth Temperament Scale 641
G. FREDERICK KUDER. Use of the International Scoring Machine for the Rapid Computation of Tables of Intercorrelations 587
GEORGE LAWTON. Mental Abilities at Senescence: A Survey of Present-Day Research 607
CLARENCE LEUBA. Has Recent Research Undermined the Evidence for Extra Sensory Perception 549

MATTHEW LUCKIESH AND FRANK K. MOSS. Effects of Leading on Readability 140
ROSS A. MCFARLAND AND CLIFFORD P. SEITZ. A Psycho-Somatic Inventory 327
GILBERT S. MACVAUGH. Regression Formula to Predict Educational Ages from Revised Alpha Scores 275
GLENN E. MITCHELL AND HAROLD E. BURTT. Psychological Factors in the Sales Interview 17
FRANK K. MOSS: See Luckiesh and Moss.
CLAUDE L. NEMZEK. The Value of the Bernreuter Personality Inventory for Direct and Differential Prediction of Academic Success as Measured by Teachers' Marks 576
JAY L. OTIS. The Prediction of Success in Power Sewing Machine Operating 350
E. DE ALTON PARTRIDGE. Psychology as an Occupation 311
DONALD G. PATERSON AND MILES A. TINKER. The Part-Whole Proportion Illusion 421
RUTH E. PERL: See Bryan and Perl.
LILLIAN PORTENIER. Aspects of Socio-Economic Status which Differentiate University Students 261
SIDNEY L. PRESSEY. A Note on the Critical Note 659
H. H. REMMERS: See Geiger, Remmers and Greenly.
MARTIN L. REYMERT AND HAROLD A. KOHN. The Mooseheart Graphic Rating Scale for Housemothers and Housefathers 288
SYDNEY ROSLOW: See Schultz and Roslow.
CHRISTIAN A. RUCKMICK. The Truth About the Lie Detector 50
ANNA LOIS SAUM. The Relative Significance of Time as a Factor In Tests of Ability at the College Level 192
NAOMI SCHILLER. Intelligence Differences in Reaction to Advertisements 169
EMMETT L. SCHOTT. Superior Intelligence in Patients with Fröhlich's Syndrome 395
RICHARD S. SCHULTZ AND SYDNEY ROSLOW. Restandardization of the A–S Reaction Study as a Personnel Form 554
CARRIE MAE SCOTT. Background and Personal Data as Factors in the Prediction of Scholastic Success in College 42
MAY V. SEAGOE. Prediction of Achievement in Elementary Algebra 493
MAY V. SEAGOE. Prediction of Achievement in Foreign Languages 632
CLIFFORD P. SEITZ. See McFarland and Seitz.
KINSLEY R. SMITH. The Relation between Age and Performance on a Repetitive Manual Task 295
MAPHEUS SMITH. An Experiment to Test the Reliability of Estimates of Use of Time 400

ROSS STAGNER, Marital Similarity in Socio-Economic Attitudes 340
R. D. STOVER: See Williamson, Stover and Fiss.
ELEROY L. STROMBERG. The Relationship of Measures of Visual Acuity and Ametropia to Reading Speed 70
DEWEY B. STUIT. A Study of the Vocational Interests of a Group of Teachers College Freshmen 527
MAY SUKOV AND E. G. WILLIAMSON. Personality Traits and Attitudes of Jewish and Non-Jewish Students 487
HOWARD R. TAYLOR: See Blankenship and Taylor.
ROBERT L. THORNDIKE. Critical Note on the Pressey Interest-Attitudes Test 657
ROBERT L. THORNDIKE. Note on "I.Q. Changes in Foster Home Children" by Emmett L. Schott 662
MILES A. TINKER: See Paterson and Tinker.
LYLE TUSSING. What Students Want from the Elementary Course in Psychology 282
VICTOR H. VOGEL: See Brown and Vogel.
F. L. WELLS, RUTH WILLIAMS, AND PHILIP FOWLER. One Hundred Superior Men 367
RUTH WILLIAMS: See Wells, Williams and Fowler.
E. G. WILLIAMSON. A Further Analysis of the Young-Estabrooks Studiousness Scale 105
E. G. WILLIAMSON, R. D. STOVER AND C. B. FISS. The Selection of Student Nurses 119
E. G. WILLIAMSON: See Sukov and Williamson.
S. JEAN WOLF. Comparison of the Standard and a "Self-Scoring" Form of the Vineland Adjustment Score Card 87
ELEANOR ZALOWITZ: See Eisenberg and Zalowitz.

Vol. XXII DECEMBER, 1938 No. 6

THE JOURNAL OF APPLIED PSYCHOLOGY

Edited by
JAMES P. PORTER, Ohio University
Athens, Ohio

TABLE OF CONTENTS

CLARENCE LEUBA. Has Recent Research Undermined the Evidence for Extra Sensory Perception 549

RICHARD S. SCHULTZ AND SYDNEY ROSLOW. Restandardization of the A–S Reaction Study as a Personnel Form 554

EDWARD B. GREENE AND VIRGINIA DAHLEM. The Michigan Vocational Preference List 558

CLAUDE L. NEMZEK. The Value of the Bernreuter Personality Inventory for Direct and Differential Prediction of Academic Success as Measured by Teachers' Marks 576

G. FREDERIC KUDER. Use of the International Scoring Machine for the Rapid Computation of Tables of Intercorrelations 587

NEAL DROUGHT. An Analysis of Eight Measures of Personality and Adjustment in Relation to Relative Scholastic Achievement 597

GEORGE LAWTON. Mental Abilities at Senescence: A Survey of Present-Day Research 607

PHILIP EISENBERG AND ELEANOR ZALOWITZ. Judging Expressive Movement: III. Judgments of Dominance-Feeling from Phonograph Records of Voice 620

MAY V. SEAGOE. Prediction of Achievement in Foreign Languages 632

BARBARA KRUGER. A Statistical Analysis of the Humm-Wadsworth Temperament Scale 641

CHARLES C. GIBBONS. A Short Scoring Method for the Link P. Q. Test 653

Notes:

ROBERT L. THORNDIKE. Critical Note on the Pressey Interest-Attitudes Test 657

SIDNEY L. PRESSEY. A Note on the Critical Note 659

ROBERT L. THORNDIKE. Note on "I.Q. Changes in Foster Home Children" by Emmett L. Schott 662

News and Notes 663

Book Reviews 666

New Books and Pamphlets 674

Index of Subjects 675

Index of Authors 683

HAS RECENT RESEARCH UNDERMINED THE EVIDENCE FOR EXTRA SENSORY PERCEPTION

CLARENCE LEUBA
Antioch College

TWO books, one technical,[1] the other popular,[2] have described the astonishing results of experiments in telepathy and clairvoyance at Duke University. It appeared from these experiments that certain people at certain times could experience what was occurring in another person's mind (telepathy), or some external event (clairvoyance), without the mediation of the senses, and with a degree of accuracy unexplainable on the basis of chance alone. This ability was called extra sensory perception.

To control chance, the Duke University investigators devised the now famous extra sensory perception cards. Each card bears one of five symbols: a cross, a rectangle, a star, a circle, or wavy lines. A subject confronted with these five cards face down, and knowing that the five different symbols were represented, had an even chance of guessing one correct out of the five. When, as was usually the case, the cards were used in packs of 25 (five crosses, five rectangles, five stars, five circles, and five wavy lines), the subject had an even chance of guessing 5 correctly out of the 25.

About one in five of the persons tested at Duke University gave results significantly above those expected on a purely chance basis. A few of these selected subjects were able to average 10, or even more, correct calls per pack of 25 cards, for many thousands of trials. On a few occasions, all 25 cards

[1] J. B. Rhine, *Extra Sensory Perception*, 1934.

[2] J. B. Rhine, *New Frontiers of the Mind*, 1937.

were called correctly. The odds against a purely chance occurrence of such phenomena were "astronomical." For good subjects, it seemed to make little difference whether the experimenter knew the order of the cards (telepathy combined with clairvoyance), or whether he did not (pure clairvoyance), or whether the experimenter just thought of the cards in some supposedly random manner (pure telepathy).

These astounding results, secured in a scientific laboratory of a great university and publicized in two books, one of which was widely read,[3] stimulated both laymen and professional psychologists, as well as the Duke University investigators themselves, to repeat these experiments and to devise safeguards against sources of error.

To prevent the subject from securing sensory cues from the cards themselves it became standard practice, in psychological laboratories at least, not to permit the subject to shuffle the cards himself or to touch them in any way.[4] During the test proper, the cards were held behind a screen. Whenever the experimenter himself knew the order of the cards, and the test was therefore one involving telepathy, he was separated from the subject by doors and walls. It is well known to psychologists that when a person thinks intently about a symbol, such as a square or a cross, the muscles involved in saying the appropriate word—square or cross—may be innervated (ideomotor action); that the person himself may be unaware of this innervation; and that persons may be influenced by sound waves of which they are not consciously aware (subliminal cues). Finally, to prevent an experimenter, under the influence of a desire for success, from misinterpreting a subject's ambiguously expressed statement of the symbol or symbols chosen, subjects were instructed to point to a picture of the symbol they had in mind.

[3] *New Frontiers of the Mind.*

[4] With certain packs it was possible, either tactually or visually, to get cues from the imprint of the symbol through the back of the card; using these cues, one psychologist was able to name all 25 cards correctly.

As these safeguards against sensory cues and recording errors were introduced at the Duke Laboratory and elsewhere, the number of successful responses, though still remaining slightly above average chance expectations, showed a marked decline. Even the best subjects reported upon in the June, 1938, issue of the *Journal of Parapsychology* averaged only between 6 and 7 correct responses per run of 25 cards.

It occurred to us that these more modest deviations from chance expectations might be due to two factors, also not extra sensory in nature:

1. The selection of unusually lucky subjects out of a large number of subjects who scored at or below the 5 hits per run expected *on the average* through chance. It is an established mathematical principle that the statistical formulae used by E.S.P. investigators apply only when all the data for all the subjects are taken into account. Nevertheless, the subjects reported upon in the literature today are chosen from a much larger, but unknown, number of subjects who may have scored at or below the five expected on the average through chance, and whose data have not been reported. It is to be expected, on a purely chance basis, that if enough subjects are tested a few will score between 6 and 7 correct for many runs through the pack of 25 cards.

2. A tendency on the part of successful experimenters to persevere at any one session, with subjects who had been lucky initially, until those subjects had had a good series of runs; and conversely, with subjects who were poor, to stop after a series of unsuccessful runs. It has been stated repeatedly in the Duke University publications that attitudes of patience and confidence were important for successful extra sensory experimentation. Just as a gambler with unlimited resources might refuse to stop at any session until he was ahead of the game, so an E.S.P. investigator, buoyed by faith, might refuse to give up with certain subjects until they were well ahead of the results expected, *on the average*, through chance.

To test empirically the influence of these two factors, it was necessary to devise an experiment in which chance alone could operate.[5] For this purpose we matched two packs of E.S.P. cards with each other, ascribing the first 10 runs to a first imaginary "subject," the second 10 runs to a second imaginary "subject," and so on until we had 87 "subjects." Those "subjects" who averaged above 5 hits per run for those first 10 runs were given a second session; the few who did well during that second session, were given a third session, and so on. With unusually good "subjects," we tended to stop after a good series of runs; with mediocre "subjects," after a poor series of runs. The 87 "subjects" made 2,682 runs (67,050 separate trials or matchings) and averaged close to the chance expectation of 5 hits per run (5.03). Seven "subjects," however, did have, at one time or another, 8 or more consecutive runs with scores which, according to the *Handbook for Testing Extra Sensory Perception* (by members of the staff of the Duke University Laboratory of Parapsychology), were clearly indicative of extra sensory powers—on the part of our cards! "Subject" B averaged 6.2 hits per run of 25 cards for his 41 runs (1025 trials; critical ratio 3.8).

This experiment provides an empirical verification of what one would expect, *a priori*, from the application of accepted mathematical principles: namely that as long as there are large numbers of people being tested for E.S.P. capacity, there are bound to be a few, who, through pure chance, will have a number of correct responses greater than the number expected *on the average* through chance. Likewise, if there are enough investigators in the field, there are bound to be a few investigators who will have an average number of correct responses for all their subjects slightly above the average chance expectation.

That both the selection of lucky subjects and the tendency to persist with certain good subjects until they were above

[5] Leuba, C., "An Experiment to Test the Role of Chance in E.S.P. Research," *Journal of Parapsychology*, 1938, 2, 217–221.

average chance expectations are effective, was shown by repeating the experiment just described under two sets of circumstances. If the experimenter was instructed to persevere with his "subjects" until above average scores were obtained, the "subjects" averaged above 5 hits per run. If the experimenter was instructed to secure a large number of "subjects," but, by giving each "subject" a predetermined number of runs (40), to prevent his own attitude from influencing the results, a few "subjects" were found who averaged distinctly above 5 hits per run for the total of 40 runs (1000 trials). In neither experiment, however, were the results as significant as in our first experiment when both factors were permitted to exercise their influence.

To summarize: The large number of correct responses obtained in the early Duke experiments and reported in two recent books, fail to appear when sensory cues and recording errors are eliminated; the more modest successes of the later experiments, under better controlled conditions, can be obtained by the selection of lucky subjects and through the influence of the experimenter's attitude in determining the point at which the testing will be discontinued.[6]

[6] There is as yet little indication that the Duke investigators and their associates have been appreciably affected by the negative evidence summarized in this article. In recent months, they claim to have demonstrated experimentally the ability of certain subjects to foretell, *before* shuffling the cards, what would be their order *after* they have been shuffled, to a degree of accuracy slightly greater than the five correct calls expected on the average through chance (pre-cognition); and the ability of certain experimenters to get reliably better results from their subjects than other experimenters, presumably because of the latter's inadequate attitudes. Both these recent developments could be ascribed, of course, to the non-E.S.P. factors mentioned in this article: the selection of unusually lucky subjects or experimenters, and the tendency on the part of successful experimenters to persevere until better than average chance results have been achieved.

RESTANDARDIZATION OF THE A–S REACTION STUDY AS A PERSONNEL FORM

RICHARD S. SCHULTZ AND SYDNEY ROSLOW
The Psychological Corporation, New York City

THE Allport Ascendance-Submission Study[1] was designed to differentiate degrees of "initiative" in social contacts. A revision[2] for business use has been made in which a number of the original items have been eliminated and in which other items have been re-written or especially prepared in order to pertain directly to business rather than to academic situations.

LIMITATION OF THE BUSINESS REVISION

The business revision has been used for some time at the Psychological Corporation in personnel research.[3,4] Its use has been restricted because the scores spread over a wide middle range, giving a flat distribution. Consequently the test does not discriminate between individuals so as to yield scores falling into a normal distribution. In an earlier investigation, the authors analyzed the results of the test administered to 544 adult men with varied employment experience who were applicants for life insurance work. The distribution shown in Table I (Business Revision) is flat between +19 and −8 (the second to the eighth decile). The distribution failed to yield a

[1] G. W. Allport. A Test for Ascendance-Submission. *Journal of Abnormal and Social Psychology*, 1928, 23, 118–196.

[2] R. O. Beckman. Ascendance-Submission Test—Revised. *The Personnel Journal*, 1933, 11, 387–392.

[3] Achilles, P. S. and Schultz, R. S. Characteristics of Life Insurance Salesmen. *The Personnel Journal*, 1934, 12, 260–263.

[4] R. S. Schultz. Standardized Tests and Statistical Procedures in Selection of Life Insurance Sales Personnel. JOURNAL OF APPLIED PSYCHOLOGY, 1936, 20, 553–566.

TABLE I

Frequency Distributions of the A-S Reaction Test

BUSINESS REVISION* Scores	f	RESTANDARDIZED SCALE** Scores	f
37–39	1	190–199	3
34–36	4	180–189	5
31–33	2	170–179	15
28–30	5	160–169	15
25–27	16	150–159	31
22–24	14	140–149	48
19–21	24	130–139	59
16–18	35	120–129	66
13–15	45	110–119	71
10–12	35	100–109	62
7– 9	50	90– 99	52
4– 6	48	80– 89	28
1– 3	48	70– 79	24
– 2– 0	46	60– 69	13
– 5– – 3	48	50– 69	7
– 8– – 6	37	40– 59	1
–11– – 9	27		
–14– –12	21	N	500
–17– –15	18	M	119.3
–20– –18	6	S.D.	29.0
–23– –21	8		
–26– –24	3		
–29– –27	1		
–32– –30	1		
–35– –33	1		
N	544		
M	4.6		
S.D.	12.5		

* Distribution of scores obtained in a previous investigation.

** Distribution of scores obtained in the present study.

clearly defined modal interval. The flatness in the middle ranges suggests that this distribution corresponds to the rectangular type of normal curve described by Kelley.[5] Others[6]

[5] T. L. Kelley. Statistical Method. Macmillan, New York, 1923.

[6] S. H. Stevens and E. F. Wonderlic. The A-S Reaction Test, A Study of Beckman's Revision. *The Personnel Journal*, 1934, 15, 222–224.

have also reported this situation in which the distribution has been flat over a wide range; namely, between +22 and +1 (the first to the fifth decile).

METHOD OF RESTANDARDIZATION

The Business Revision of the A–S Reaction Study was administered to a new group of 500 adult men, applicants for sales positions. These men have had a wide variety of employment experience in all types of occupations and business. Their ages ranged from 21 to 45, and their education from sixth grade to college graduate. The men were chosen at random from six cities; New York, Philadelphia, Chicago, Syracuse, Boston and Columbus. The tests were scored according to the key supplied with an earlier revision.[2] On the basis of these scores, the upper 27 per cent and the lower 27 per cent were selected. An item analysis was made using the upper and lower 27 per cent method attributed to Kelley.[7] A new scoring system was derived and the 500 tests were rescored. The distribution of the revised scores is shown in Table I.

An inspection of the table indicates that this distribution is decidedly more normal than the distribution of the Business Revision discussed above. The histogram for this frequency distribution and the theoretical normal curve fitted to these data have been omitted here. The goodness of fit was determined by the chi-square function. The value of P was found to be .98. Holzinger[8] states that unless P is .2 or more the fit should not be regarded as good. The goodness of fit in this case is unusual. Since the investigation is concerned not with the question of sampling but with the construction of a measuring device to yield scores which distribute normally, the value of P indicates that the latter aim has been achieved.

[7] J. A. Long and P. Sandiford. The Validation of Test Items. Bulletin No. 3, 1935. Department of Educational Research, University of Toronto.

[8] K. J. Holzinger. Statistical Methods for Students in Education. Ginn & Co., Boston, 1928, p. 247.

GENERAL SUMMARY

This study is the first step in the restandardization of the Business Revision of the A–S Reaction Study. The aim is to develop a personality test for adults to be used in personnel administration. The Business Revision has been limited in use because it spread the scores over a large range tending to give a rectangular distribution. The restandardization gave a more normal distribution. The next step in the research program now in progress is to accumulate data on other adult groups and to construct norms.*

* Scoring procedure, reliability, and validity will be published subsequently.

THE MICHIGAN VOCATIONAL PREFERENCE LIST

EDWARD B. GREENE AND VIRGINIA DAHLEM
University of Michigan

THE PROBLEM

TWO types of Vocational Preference Lists are to be found. One type, called Form I in this report, has the items arranged in alphabetical order. It is illustrated by Strong's (1933) Vocational Interest Blank, Thurstone's (1935) Vocational Interest Schedule, Manson's (1931) Occupational Interests for Women, and many others. The other type, called Form II, has the items arranged by divisions which are designed to be homogeneous. This is illustrated by the Michigan Vocational Preference List, Table 4. Form II has divisional headings which may influence the responses to the subheadings by establishing a mental or emotional "set." Form I has no divisional headings, hence the "set" for each response must vary, depending somewhat on the previous item. This requires a shift from the consideration of one occupation to another which is often far removed from it. Thus in Thurstone's Vocational Interest Schedule the first four items are "actor, advertiser, architect, and army officer."

A serious obstacle to the widespread use of Form I is the expense of scoring. The scoring must be done with a different stencil for each occupation or factor. Strong (1933) prepared twenty-two stencils and Thurstone (1935), seven. From these a profile showing one's relative likes and dislikes is prepared. To eliminate that difficulty Form II was designed so that a score could be obtained for each division by finding the proportion of items marked L (Like). This form may be quickly scored without the use of stencils. It is expected that when Forms

I and II have the same items they will give profiles which are substantially the same. Reasons for this expectation are:

1. The analyses of Strong's (1934) results showed that there was considerable similarity between a number of the scoring keys. Dwyer (1937) has taken these keys and found that the present 22 Strong patterns of occupational preferences and 3 patterns of maturity, studiousness, and masculinity can be calculated with remarkable accuracy from only five stencils: those for, Physicist, Teacher, Life-insurance Salesman, Journalist, and Office Worker.

2. Thurstone's (1935) analysis showed that certain items were more highly saturated than others with particular factors. Hence a selection of items could be made which would represent each factor fairly adequately.

3. Studies of the permanence cf interest by Franklin (1926), Fryer (1924), and Strong (1931) showed that while one's choices of a particular occupation varied considerably, one's main division of interest tended to remain fairly constant.

In order to ascertain the effects of grouping, two forms were devised having the same items and directions and no duplicate items. In Form I the items were arranged alphabetically. In Form II the items were grouped under eleven headings.

THE CHOICE OF ITEMS

A list of items was desired which would include the names of the most important occupations and professions for both men and women, and have as little ambiguity as possible. A group of seven judges, graduate students in psychology, cooperated in the following procedure:

1. The 283 most important occupations and professions in the 1930 census were listed on the basis of frequency and popularity of choice as shown by previous check lists.

2. Occupational names which seemed ambiguous were discarded. Thus, in place of "florist," which refers either to one who sells flowers or one who grows flowers, "storekeeper" and "nurseryman" were chosen. In place of "banker," "bank clerk" and "bank official" were chosen.

3. Items which had nearly the same meaning were studied and in some cases one was eliminated.

From this work 181 items were finally selected.

COMPARISONS WITH OTHER LISTS OF OCCUPATIONS

In order to show the amounts of overlapping between various vocational lists, Table 1 was prepared. This shows that all the lists compared have approximately 70 per cent of items identical with, and 14 per cent of items similar to the list prepared in this study. Another 11 per cent of items from these lists were more general than the Michigan List including

TABLE 1

A Comparison of Items in Various Lists with the Present Study

MICHIGAN FORM	STRONG				THURSTONE		CLEETON				MANSON	
	Men		Women		Men		Men		Women		Women	
	N	%	N	%	N	%	N	%	N	%	N	%
Identical with	69	69	87	68	52	70	56	77	91	74	111	70
Similar to	14	19	24	19	10	13	8	11	12	10	26	16
More general	14	14	12	9	8	11	8	11	15	12	8	5
More rare	3	3	5	4	4	5	1	1	5	4	15	9
Total	100	100	128	100	74	100	73	100	123	100	160	100

reporter, manufacturer, civil service employee, college professor, factory worker, author of technical book, public speaker, farmer, scientific research worker, magazine writer, etc. Approximately 5 per cent of items were found to be more rare or not to constitute a well defined occupation. These included auctioneer, dean of women, minister's assistant, taxidermist, governor of a state, labor arbitrator, notary public, etc.

It is evident from this comparison that the Michigan List includes the important occupations for men and women which are listed in the other check lists studied. The Michigan list differs from the others in having a larger number of trade names and a smaller proportion of general items. These general items are, however, well represented in the present study in the form of more specific occupational names. It was hoped that this greater specificity would lead to more precision in

ratings and hence a better analysis of interest, particularly among persons in semi-skilled trades, where Paterson and Darley (1936) found that the Strong scale was not discriminative.

THE CHOICE OF DIVISIONS

An inspection of the items selected showed very wide differences in prerequisite training. To help evaluate this factor a study was made of technical and professional school catalogues and of Freyer's 1922 report. The average years of training needed was then placed after each item. There was considerable disagreement among the judges concerning some of the items, due to the fact that in some occupations the standards of training are variable. The numbers selected represent the approximate years of training needed for average success by the average person in the occupation in the Middle West.

From a study of these figures, two large groups appeared: A. of occupations which needed more than 12th grade training or ability, and B. of occupations which generally needed less than 13 years of average training. These two groups served as a rough basis of further classification because the 12th grade is a fairly definite educational level, and there is a large gap between 12th grade work and further study.

A further classification of occupations according to common occupational objectives and techniques was made by each judge using a variation of a paired comparison method. This resulted in six divisions in Group A. and six corresponding divisions in Group B. as shown in Table 2.

There was little disagreement among the judges concerning these main divisions. It was felt by some that the sixth division, Government and Legal, could be logically divided between the Social and Commercial divisions. However, it was found to represent a large field with a special vocabulary and techniques, so it was included. Some of the judges also felt that a division of occupations which required special language training was needed. After considerable debate this motion was lost, because it appeared that the language training was, with few

TABLE 2

Subjective Classification of Occupations

A. ABOVE 12TH GRADE	B. BELOW 13TH GRADE
1. *Social Professions*	7. *Recreation*
Teaching	Sports
Ministry	Entertaining
Personnel	8. *Office Techniques*
Social Work	Clerical
2. *Commercial*	Accounting
Business	Telegraphy
Manufacture	9. *Mechanical Trades*
Sales	Iron Work
Economics	Wood Work
3. *Physical Sciences*	Mason Work
Chemical	Transportation
Physical	Electrical
Engineering	10. *Agricultural Trades*
4. *Biological Sciences*	Food Growing
Zoology	Forester
Nutrition	Horticulture
Medicine	11. *Personal Service*
5. *Fine Arts*	Dressmaking, Tailor
Plastic, Graphic	Cleaning, Laundry
Music	Hair Dressing
Drama	Food Preparation and Service
Literature	12. *Police*
6. *Government and Legal*	Local, State, Federal
Legislative	Soldier
Executive	Sailor
Judicial	
Military	

exceptions, subservient to writing about a field already included, or to the teaching of language. Such occupations as foreign correspondent, publicity agent, art critic, etc., were therefore placed in government, commerce, and art divisions, respectively. A similar motion for a mathematical division was lost for similar reasons. It was agreed that the classification might be changed on the basis of future findings. It was ad-

mitted that the classification emphasized practical vocational training, rather than functional or basic capacities or satisfactions.

The choice of divisions was further aided by inspecting the analyses of Strong (1934), Thurstone (1935), and Cleeton (1935) shown in Table 3. The divisions used in this study correspond very closely with factor analyses of Thurstone and Strong in occupations above the 12th grade and fairly well with Cleeton's classification throughout. The differences in classification are due to variations in approach to the selection of items. Only from more careful definitions and factor analyses of large groups of items can such classification be made more effectively.

TABLE 3

Classification of Occupations Compared

PRESENT STUDY	THURSTONE 1935	STRONG 1934	CLEETON 1935
1. Social and Ed. Service	Descriptive	People (Maturity)	Teaching Ministry
2. Commercial	Commercial	Business	Sales Accounting
3. Physical Science	Physical Science	Science (Intellect)	Physical Science
4. Biological Science	Biological Science	Science	Biological Science Nursing
5. Art	(Art)	————	Artistic
6. Government Legal	Legal	————	————
7. Recreation	Athletics	————	————
8. Office Techniques	————	————	Office and Clerical
9. Mechanical Trades	————	————	Mechanical Trades
10. Agricultural Trades	————	————	————
11. Personal Service	————	————	Personal Service
12. ————	Academic	Language (Femininity)	Literary Legal

The actual placing of an item in a division resulted in very few disagreements. These appeared, in most cases, in making decisions between the A and B groups. Thus, a few items in Divisions 7 to 11 inclusive require more than 12th grade ability. These items were placed as they appear in Table 4, because they seemed to have more in common with the techniques of the B group. The 6th and 12th divisions were combined because so few items were obtained for each. A detailed factor analysis is planned for these items to see to what extent each item may be dependent upon common factors, among various age, sex, and occupational groups.

TABLE 4

Michigan Vocational Preference List: Form II

Name _____ Age _____ Sex M. F. Date _____

Directions: If you would like the work in an occupation listed below, circle the (L) in front of it. If you would dislike it, circle the (D). If you don't know or if you are neutral, circle the (?). Please check all items. Work rapidly. Salary and working conditions are to be checked on another sheet, so do not allow them to influence your choices here.

1. Social and Educational Service

1. L ? D Boy or girl scout leader (12)
2. L ? D Care of small children (12)
3. L ? D Parole, probation officer (14)
4. L ? D Kindergarten teacher (15)
5. L ? D Missionary; home, foreign (15)
6. L ? D Grade school teacher (16)
7. L ? D Visiting teacher (16)
8. L ? D Civics teacher (16)
9. L ? D Language teacher (16)
10. L ? D Clergyman or priest (16)
11. L ? D Summer camp director (16)
12. L ? D Personnel director (16)
13. L ? D Columnist, human interest (16)
14. L ? D Social service worker (17)
15. L ? D Librarian (17)
16. L ? D Psychologist, clinical (18)
17. L ? D Sociologist (19)
18. L Other _____

TABLE 4.—(*Continued*)

2. Commercial

1.	L	?	D	Clerk, drug store (9)
2.	L	?	D	Clerk, retail sales (9)
3.	L	?	D	Demonstrator (10)
4.	L	?	D	Manager, movie house
5.	L	?	D	Book agent
6.	L	?	D	Commercial traveler (13)
7.	L	?	D	Merchant, retail (13)
8.	L	?	D	Store keeper, head (14)
9.	L	?	D	Real estate agent (14)
10.	L	?	D	Personal Finance agent (14)
11.	L	?	D	Radio announcer (14)
12.	L	?	D	Salesman, automobile (15)
13.	L	?	D	Merchant, wholesale (16)
14.	L	?	D	Insurance agent (16)
15.	L	?	D	Buyer, Department store (16)
16.	L	?	D	Publicity agent (16)
17.	L	?	D	Press agent (16)
18.	L	?	D	Stock broker (16)
19.	L	?	D	Sales manager (16)
20.	L	?	D	Hotel manager (17)
21.	L	?	D	Bank officer (18)
22.	L			Other .....

3. Physical Science

1.	L	?	D	Mathematics teacher (16)
2.	L	?	D	Surveyor (16)
3.	L	?	D	Building Contractor (16)
4.	L	?	D	Factory manager (16)
5.	L	?	D	Aviator (16)
6.	L	?	D	Chemist (16)
7.	L	?	D	Architect (16)
8.	L	?	D	Mechanical engineer (16)
9.	L	?	D	Electrical engineer (16)
10.	L	?	D	Physics teacher (17)
11.	L	?	D	Explorer, scientific (17)
12.	L	?	D	Inventor (18)
13.	L	?	D	Geologist (18)
14.	L	?	D	Astronomer (18)
15.	L			Other .....

TABLE 4.—(*Continued*)

4. Biological Science

1.	L	?	D	Dental hygienist (13)
2.	L	?	D	Osteopath (15)
3.	L	?	D	Physiotherapist (15)
4.	L	?	D	Nurse, graduate (16)
5.	L	?	D	Dietitian (16)
6.	L	?	D	Pharmacist (16)
7.	L	?	D	Hospital Laboratory Technician (16)
8.	L	?	D	Biology teacher (17)
9.	L	?	D	Botany teacher (17)
10.	L	?	D	Dentist (18)
11.	L	?	D	Bacteriologist (18)
12.	L	?	D	Physician (20)
13.	L	?	D	Physiologist (20)
14.	L	?	D	Surgeon (20)
15.	L	?	D	Oculist (20)
16.	L	?	D	Psychiatrist (20)
17.	L			Other .....

5. Arts

1.	L	?	D	Artists model (9)
2.	L	?	D	Dancer (12)
3.	L	?	D	Movie star (12)
4.	L	?	D	Illustrator (13)
5.	L	?	D	Actor or actress (14)
6.	L	?	D	Scenario writer (14)
7.	L	?	D	Modiste (15)
8.	L	?	D	Musician (16)
9.	L	?	D	Singer (16)
10.	L	?	D	Sculptor (16)
11.	L	?	D	Painter, artist (16)
12.	L	?	D	Interior decorator (16)
13.	L	?	D	Novelist (16)
14.	L	?	D	Orchestra conductor (17)
15.	L	?	D	Poet (17)
16.	L	?	D	Music teacher (17)
17.	L	?	D	Art teacher (17)
18.	L	?	D	Art critic (18)
19.	L	?	D	Playwright (18)
20.	L			Other .....

TABLE 4.—(*Continued*)

6. Government and Legal

1.	L	↑	D	Soldier or sailor (8)
2.	L	↑	D	Police officer or sheriff (10)
3.	L	↑	D	Detective (12)
4.	L	↑	D	State patrolman (12)
5.	L	↑	D	Political organizer (15)
6.	L	↑	D	Army officer (16)
7.	L	↑	D	Government investigator (criminal) (16)
8.	L	↑	D	Congressman (16)
9.	L	↑	D	Interpreter (16)
10.	L	↑	D	Foreign correspondent (16)
11.	L	↑	D	Feature writer, political (17)
12.	L	↑	D	Diplomatic service (18)
13.	L	↑	D	Lawyer, criminal (19)
14.	L	↑	D	Lawyer, corporation (19)
15.	L	↑	D	Justice or Judge (19)
16.	L			Other _____

7. Sports

1.	L	↑	D	Baseball player (8)
2.	L	↑	D	Professional athlete (8)
3.	L	↑	D	Golf professional (10)
4.	L	↑	D	Swimming instructor (10)
5.	L	↑	D	Billiard expert (10)
6.	L	↑	D	Bowling expert (10)
7.	L	↑	D	Playground supervisor (14)
8.	L	↑	D	Physical Education teacher (16)
9.	L	↑	D	Gymnasium teacher (16)
10.	L	↑	D	Athletic coach (16)
11.	L			Other _____

8. Office

1.	L	↑	D	Telephone switchboard operator (9)
2.	L	↑	D	Filing clerk (9)
3.	L	↑	D	Cashier, store (9)
4.	L	↑	D	Typist (10)
5.	L	↑	D	Mail clerk (10)
6.	L	↑	D	Bookkeeper (11)
7.	L	↑	D	Ticket agent, railway (12)
8.	L	↑	D	Postmaster or mistress (12)
9.	L	↑	D	Stenographer (12)

TABLE 4.—(*Continued*)

10.	L	↑	D	Telegraph operator (12)
11.	L	↑	D	Bank clerk (12)
12.	L	↑	D	Secretary, private (12)
13.	L	↑	D	City or county clerk (12)
14.	L	↑	D	Linotype operator (12)
15.	L	↑	D	Statistical worker (13)
16.	L	↑	D	Office manager (16)
17.	L	↑	D	Certified Public Accountant (19)
18.	L			Other .....

9. Mechanical Trades

1.	L	↑	D	Janitor or watchman (7)
2.	L	↑	D	Chauffeur or taxi driver (7)
3.	L	↑	D	Machine tender (7)
4.	L	↑	D	House painter (9)
5.	L	↑	D	Carpenter (10)
6.	L	↑	D	Mason (10)
7.	L	↑	D	Plumber (10)
8.	L	↑	D	Welder (10)
9.	L	↑	D	Stationary engineer (10)
10.	L	↑	D	Machinist (12)
11.	L	↑	D	Sheet metal worker (12)
12.	L	↑	D	Electrician (12)
13.	L	↑	D	Garage mechanic (12)
14.	L	↑	D	Bus driver (12)
15.	L	↑	D	Fire Department member (12)
16.	L	↑	D	Telephone lineman (12)
17.	L	↑	D	Radio operator (13)
18.	L	↑	D	Locomotive fireman (13)
19.	L	↑	D	Printer, machine (13)
20.	L	↑	D	Instrument maker (14)
21.	L	↑	D	Draftsman (15)
22.	L	↑	D	Foreman, shop (15)
23.	L			Other .....

10. Agriculture

1.	L	↑	D	Rancher, cattle (8)
2.	L	↑	D	Farmer, general (10)
3.	L	↑	D	Poultry raiser (10)
4.	L	↑	D	Forester (10)
5.	L	↑	D	Truck farmer (10)

TABLE 4.—(*Concluded*)

6.	L	?	D	Tobacco farmer (10)
7.	L	?	D	Cotton grower (10)
8.	L	?	D	Dairyman (12)
9.	L	?	D	Fruit grower (12)
10.	L	?	D	Grain grower (12)
11.	L	?	D	Nurseryman (14)
12.	L	?	D	Agricultural experimenter (16)
13.	L	?	D	Government Agricultural agent (16)
14.	L			Other .....

11. Personal Service

1.	L	?	D	Porter (4)
2.	L	?	D	Domestic helper (7)
3.	L	?	D	Waiter or waitress (7)
4.	L	?	D	Laundry worker (7)
5.	L	?	D	Manicurist (7)
6.	L	?	D	Cook (8)
7.	L	?	D	Barber (8)
8.	L	?	D	Corsetiere (9)
9.	L	?	D	Housekeeper (9)
10.	L	?	D	Tailor (9)
11.	L	?	D	Hairdresser (10)
12.	L	?	D	Cleaner and dyer (11)
13.	L	?	D	Milliner (11)
14.	L	?	D	Dressmaker (12)
15.	L	?	D	Domestic science teacher (16)
16.	L	?	D	Caterer or chef (16)
17.	L	?	D	Hotel hostess (16)
18.	L			Other .....

Considering all the items of this list please indicate your first three choices of occupations by placing a (1) in front of your first choice, a (2) in front of the second, and a (3) in front of the third.

CHANGES DUE TO GROUPING OF ITEMS

The results of grouping items as in Form II, will be discussed under three headings:

1. Do Forms I and II cause different distributions of answers?

2. Which Form is more reliable?

3. Is there a halo effect due to grouping homogeneous items together?

In order to answer these questions the cooperation was secured of 211 college sophomores and juniors, most of whom were enrolled in two psychology courses in the Literary College of the University of Michigan. Their ages and chosen occupations are shown in Table 5.

TABLE 5

Data on Subjects

CHOSEN OCCUPATION	MEN	WOMEN
N	142	69
MEAN AGE	18–7	18–4
S.D.	1–2	1–2
1. Social and Educational	17	26
2. Commercial	74	18
3. Physical Science	9	1
4. Biological Science	21	9
5. Fine Arts	7	10
6. Government Legal	14	1
7. Undecided	0	4

This shows that a majority of the men were planning to go into business, and a third of the women into social and educational work. These choices should be kept in mind in interpreting the results which follow. Each subject completed each form twice, once at intervals of one day, and again, one week later. The forms were rotated to avoid practice effects.

Results

1. *Distribution of Answers.* In order to answer the first problem, "Were there significant differences between the two forms in the distribution of answers?" a tabulation was made of the frequency of the ratings of Like (L) Neutral (?) and Dislike (D) for each form. The results shown in Table 6 indicate that there were no significant differences of this sort either between the two forms or between the two trials of one form. There were small but consistent sex differences. The

men showed a slightly greater proportion of "likes" than the women, and a slightly smaller proportion of "dislikes" than the women. The men and women showed nearly the same proportion of indifferent or neutral answers.

TABLE 6

Percentage of Likes, Dislikes and Neutral Answers

	MALE			FEMALE		
	L	?	D	L	?	D
Form I (Alphabetical)						
1st trial	17.6	15.3	67.1	15.1	17.3	77.6
2nd trial	14.9	11.6	74.5	11.7	16.4	71.9
Form II (Grouped)						
1st trial	18.3	14.4	67.3	14.1	14.0	71.9
2nd trial	16.4	15.6	68.0	12.3	15.1	72.6

2. *The Reliability of Each Form.* The correlations between two trials of the same form for each person were secured by an unlike sign method adapted to three variables, which has been used throughout. Form I proved to be less consistent (average $r = .76$) than Form II ($r = .89$). The difference of .13 is seven times its own P.E. It must be concluded that the effect of Form II was to reduce considerably the random changes between trials. In other words, when the 181 items were grouped under these eleven headings the students were more consistent in their ratings than when the items were in alphabetical order.

3. *The Correlations between Forms.* The correlations between Forms I and II were .84 and .81 on the first and second trials respectively. The fact that the correlations between Forms I and II were slightly higher than the self-correlation of Form I needs an explanation. This is probably because trials of Forms I and II were given on adjacent days, while an interval of one week intervened between two trials of the same Form.

4. *Halo Effects.* If halo effects were present when the items were grouped they could be detected by comparing random changes with the experimental changes. The average random

changes should theoretically be zero. A halo effect would be indicated by a tendency to shift in one direction more than the other. In this investigation random changes were measured by tabulating the amount and direction of the changes in ratings which occurred between two trials of Form I, and two trials of Form II. These random changes were averaged.

The average experimental changes were indicated by the amount and direction of changes in ratings between Form I (alphabetical) and Form II (grouped). The calculation of directional amount of shifting was made for each student as follows:[1]

1. The letters circled in each column were summed for each trial of each form.
2. To secure an index of random shift, the sums of the second trials were subtracted from the sums of the first for each form. These differences for Form I and Form II were added. (As a check, the plus differences must always equal the minus differences.)
3. An index which showed the direction and amount of shift was found by multiplying the L's by plus one, the I's by zero,

TABLE 7

The Calculation of a Halo Effect from one Subject in One Division. (*Commercial*)

	COLUMN TOTALS								
	Trial 1			Trial 2			Diff. 1–2 (Random)		
	L	?	D	L	?	D	L	?	D
Form I	7	3	9	5	6	8	+2	−3	+1
Form II	5	4	10	3	2	14	+2	+2	−4
Diff. I–II (Experimental	+2	−1	− 1	+2	+4	− 6			
Sums of Diffs.	+4	+3	−7				+4	−1	−3
Indices	11 L Experimental						+7 L Random		

Note: This case shows four more changes toward the L rating in the experimental shifting than in the random shifting.

[1] We are indebted to Paul S. Dwyer of the Bureau of Education Investigations for aid in planning these calculations and for checking the correlational methods used.

and the D's by minus one, and adding. Plus values then indicate a shift toward the Like (L), and minus toward the Dislike (D).

4. To secure an index of experimental shift, the sums of Form II were subtracted from the sums of Form I. The procedure thereafter was identical with that described in 3 above. An example of the calculation of indices of shift is shown in Table 7.

These calculations were completed for each person on each of the first three divisions: Social and Educational Service, Commerce, and Physical Sciences. The cost of completing these computations for all the divisions proved prohibitive. The results on the first two divisions are considered to be more significant than the others because they were the fields most frequently chosen by these students. The average random shifts and experimental shifts are shown in Table 8. This table shows

TABLE 8

Random vs. Experimental Shifting ($N = 211$)

	AV.	S.D.	DIFF./S.D.$_D$
1. *Social and Ed. Service*			
Random shift	.68L	5.6	
Experimental shift	1.30L	4.2	
Difference	.72L		1.91
2. *Commercial*			
Random shift	.25L	5.4	
Experimental shift	1.37L	5.2	
Difference	1.12L		1.96
3. *Physical Science*			
Random shift	.12L	3.0	
Experimental shift	.00	2.9	
Difference	.12L		.37

that all tendencies to shift were always positive and also near zero. The first two divisions showed a slightly larger experimental than random shift toward the "Likes" but differences were so small as to be of little significance. We must conclude, then, that a halo effect was present in the first division which caused an average shifting of .72 of one item of the nineteen, from a rating of Dislike (D) to Indifference (?), or from a

rating of Indifference (?) to Like (L). In the commercial division the halo effect was a little larger, 1.12 items of the twenty-one were shifted toward the "Likes" by the average student. In the case of the Physical Sciences, the shift was in the opposite direction for .12 of an item among the fourteen was rated more toward the "Dislikes" than in the random shifting of ratings. The last case is a nearly chance variation. Whether or not these shifts were due to cognitive or affective groupings cannot be determined from this data.

SUMMARY AND CONCLUSIONS

One hundred and eighty-one names of occupations and professions for men and women were selected from the 1930 United States Census on the basis of number employed and the type of work. These names were arranged alphabetically in Form I. In Form II, the same items were grouped in twelve divisions. The first six divisions included professions which usually require more than 12th grade training: 1. Social and Educational Service, 2. Commerce, 3. Physical Science, 4. Biological Science, 5. Fine Arts, 6. Government and Legal. The last six divisions included occupations which usually require less than 13th grade training: 7. Recreation, 8. Office Work, 9. Mechanical Trades, 10. Agriculture, 11. Personal Service, 12. Police.

Two hundred eleven college students completed two trials each of Form I and II. The forms were rotated so as to avoid practice effects.

The results showed no significant differences in the distribution of ratings between Forms I and II. Form II (grouped) was answered more consistently on the two trials ($r_{11} = .89$) than Form I (.76). The correlations between the two forms were approximately .825. There were very slight halo effects, shown by a tendency to "like" the occupations in the Social and Educational Service, and in Commerce, more in the grouped than in the alphabetical form. A very slight tendency to "like" the occupations in the Physical Sciences less in the grouped than the alphabetical form was noted.

From these results it must be concluded that the grouping of these items under eleven headings improved the reliability of the form, but resulted in no significant changes in distribution of ratings, and caused no marked halo effects.

Other investigations are needed to show whether these results will obtain from other age groups or from other occupational groups.

REFERENCES

CLEETON, G. U. 1935. Occupational Adjustment. Pittsburgh Personnel Association.

DWYER, PAUL S. 1937. A Study of Relationships between Strong's Occupational Interests. Unpublished. U. of Michigan Bureau of Ed. Investigation.

FRANKLIN, E. E. 1926. The Permanence of Vocational Interests after Three Years. Sch. & Soc. 23, 438–440.

FRYER, D. 1922. Occupational Intelligence Standards. Sch. & Soc. 16, 273–277.

FRYER, D. 1924. The Significance of Interest for Vocational Prognosis. Ment. Hyg. 8, 466–505.

MANSON, GRACE E. 1931. Occupational Interests and Personality Requirements of Women in Business and Professions. Michigan Business Studies 8, 3.

PATERSON, D. G. and DARLEY, J. G. 1936. Men, Women, and Jobs. Univ. of Minnesota Press, pp. 145.

STRONG, E. K. 1931. Changes of Interest with Age. Stanford Univ. Press.

STRONG, E. K. 1933. Vocational Interest Blank. Stanford Univ. Press.

STRONG, E. K. 1934. Classification of Occupations by Interests. Personnel Journal 12, 301–313.

THURSTONE, L. L. 1935. Vocational Interest Schedule. Chicago Univ. Press.

THE VALUE OF THE BERNREUTER PERSONALITY INVENTORY FOR DIRECT AND DIFFERENTIAL PREDICTION OF ACADEMIC SUCCESS AS MEASURED BY TEACHERS' MARKS*

CLAUDE L. NEMZEK
University of Detroit

DURING recent years numerous investigations have been made to measure non-intellectual factors conditioning academic success. These investigations have closely followed the development and construction of new measuring instruments. The present study contains an analysis of the Bernreuter Personality Inventory and attempts to reveal any value it may have for direct and differential prediction of academic success as measured by teachers' marks.

The Bernreuter Personality Inventory has been administered to sophomore students at the University High School of the University of Minnesota in April of each year for a number of years. Adequate data were available for 92 boys and 99 girls on the following eight variables:

(1) intelligence quotients,
(2) B1–N scores,
(3) B2–S scores,
(4) B4–D scores,
(5) honor point averages in mathematics,
(6) honor point averages in English,
(7) honor point averages in history and social science,
(8) honor point averages in languages.

* From a thesis submitted to the Graduate Faculty of the University of Minnesota in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

The measure of intelligence used in this study was based upon the results of five group intelligence tests. The tests employed were Army Alpha 8, Pressey Senior Classification, Haggerty Delta 2, Terman Group Test, Form A, and Miller's Mental Ability Test, Form A. Intelligence quotients were computed from the results of each test for each individual. The author's manuals for the respective tests were followed as closely as possible in administering and scoring, and, for all cases except that of the Pressey Test, in computing the intelligence quotients. In this instance, where the author's norms proved inadequate for children who made unusually high scores, the difficulty was resolved by extrapolation. The intelligence quotients were in all instances converted into Stanford-Binet equivalents by means of the method proposed by Miller.[1] Of the five intelligence quotients, the middle value was chosen as the measure to be used for each individual.

Marks at University High School are given in the form of letter ratings. For the present study the letter ratings were converted into honor point averages. Each quarter hour mark of A was given three honor points; each quarter hour mark of B, two honor points; each quarter hour mark of C, one honor point; each quarter hour mark of D, no honor points; and each quarter hour mark of F, minus one honor point. Then the total number of honor points in each of the four subject-matter fields was divided by the total number of quarter hours of marks involved in the respective subject-matter fields in order to obtain the four honor point averages.

Only those cases were retained for study which had two or more years of courses in each of the four subject-matter fields; namely, mathematics, English, history or social science, and languages. All of the students had completed at least the junior year of high school; most of them had been graduated from the University High School when the present analyses were made.

[1] W. S. Miller, "The Variation and Significance of Intelligence Quotients Obtained from Group Tests," *Journal of Educational Psychology*, 15: 359–366, 1924.

The zero order Pearson product-moment coefficients of correlation were calculated between the several pairs of the eight variables for the boys and girls separately. Tables I and II

TABLE I

Pearson Product-Moment Coefficients of Correlation with Their Probable Errors Based upon (1) Intelligence Quotients, (2) B1–N Scores, (3) B2–S Scores, (4) B4–D Scores, (5) Honor Point Averages in Mathematics, (6) Honor Point Averages in English, (7) Honor Point Averages in History and Social Science, and (8) Honor Point Averages in Languages of 92 Boys.

VARIABLES	2	3	4	5	6	7	8
1	.102	.034	−.088	.438	.460	.538	.458
	.070	.070	.070	.057	.055	.050	.056
2		−.420	−.752	−.113	−.064	−.027	−.0003
		.058	.031	.069	.070	.070	.070
3			.498	.197	.166	.039	.034
			.053	.068	.068	.070	.070
4				.037	.089	.002	−.057
				.070	.070	.070	.070
5					.630	.670	.641
					.042	.039	.041
6						.803	.781
						.025	.034
7							.766
							.029

include the zero order coefficients for boys and girls, respectively. Table III presents the means and standard deviations for both sexes.

A perusal of Tables I and II suffices to show that the Bernreuter Personality Inventory has practically no value for purposes of predicting academic success as measured by teachers' marks. The data indicate that the mental functions measured by the Bernreuter Personality Inventory have little, if anything, in common with the mental functions measured by teachers' marks. In the case of the boys, the B1–N scores correlate $-.113 \pm .069$ with honor point averages in mathematics; $-.064 \pm .070$, with honor point averages in English;

TABLE II

Pearson Product-Moment Coefficients of Correlation with Their Probable Errors Based upon (1) Intelligence Quotients, (2) B1–N Scores, (3) B2–S Scores, (4) B4–D Scores, (5) Honor Point Averages in Mathematics, (6) Honor Point Averages in English, (7) Honor Point Averages in History and Social Science, and (8) Honor Point Averages in Languages of 99 Girls.

VARIABLES	2	3	4	5	6	7	8
1	−.114	.153	.095	.583	.602	.486	.571
	.067	.066	.067	.045	.043	.052	.046
2		−.243	−.658	−.134	−.093	−.106	−.045
		.064	.038	.067	.067	.067	.068
3			.432	.087	.081	.131	.056
			.055	.067	.067	.067	.068
4				.033	.073	.111	−.064
				.068	.067	.067	.068
5					.786	.721	.715
					.026	.033	.033
6						.845	.787
						.019	.026
7							.708
							.034

TABLE III

Means and Standard Deviations of (1) Intelligence Quotients, (2) B1–N Scores, (3) B2–S Scores, (4) B4–D Scores, (5) Honor Point Averages in Mathematics, (6) Honor Point Averages in English, (7) Honor Point Averages in History and Social Science, and (8) Honor Point Averages in Languages for 92 Boys and 99 Girls.

VARIABLES	BOYS (N = 92)		GIRLS (N = 99)	
	Mean	S.D.	Mean	S.D.
1	119.15	11.80	117.45	12.20
2	− 92.00	71.50	− 76.25	83.00
3	14.50	45.75	− 20.25	52.00
4	33.75	61.00	21.25	59.00
5	1.130	.695	1.055	.735
6	1.055	.640	1.410	.650
7	1.295	.735	1.225	.795
8	1.105	.785	1.400	.855

– .027 ± .070, with honor point averages in history and social science; and – .0003 ± .070, with honor point averages in languages. Only one of these coefficients is even larger than its probable error, and in that instance, – .113 ± .069, the relationship is too small to be of any statistical significance.

Further study of Table I shows that B2–S scores correlate .197 ± .068 with honor point averages in mathematics; .166 ± .068, with honor point averages in English; .039 ± .070, with honor point averages in history and social sciences; and .034 ± .070, with honor point averages in languages. The latter two coefficients are of a smaller magnitude than their probable errors; the former two, while not statistically significant, show some degree of positive relationship. When intelligence is held constant by means of partial correlation, the relationship between the B2–S scores and the honor point averages in mathematics becomes .203 ± .067. This coefficient does not meet the usual criterion for statistical significance, but it is 3.03 times its probable error. One must conclude, however, that it has negligible value for prediction purposes. Similarly, the correlation between B2–S scores and honor point averages in English becomes .169 ± .068 when intelligence is held constant by means of the partial correlation technique. This coefficient is not statistically significant, being 2.49 times its probable error. Moreover, it has negligible value for purposes of prediction.

The B4–D scores correlate .037 ± .070 with honor point averages in mathematics; .089 ± .070, with honor point averages in English; .002 ± .070, with honor point averages in history and social science; and – .057 ± .070, with honor point averages in languages. Only one of these coefficients is even larger than its own probable error.

Similar data for girls indicate that the B1–N scores correlate – .134 ± .067 with honor point averages in mathematics; – .093 ± .067, with honor point averages in English; – .106 ± .067, with honor point averages in history and social science; and – .045 ± .068, with honor point averages in languages.

Not one of these coefficients meets the usual criterion of statistical significance.

The data show that the B2–S scores correlate .087 ± .067 with honor point averages in mathematics; .081 ± .067, with honor point averages in English; .131 ± .067, with honor point averages in history and social science; and .056 ± .068, with honor point averages in languages. No statistical significance may be attached to any of these coefficients.

The B4–D scores correlate .033 ± .068 with honor point averages in mathematics; .073 ± .067, with honor point averages in English; .111 ± .067, with honor point averages in history and social science; and – .064 ± .068, with honor point averages in languages. These coefficients fail to reveal any significant relationships.

The present data justify the same conclusion given by Finch and Nemzek[2] in a previous study in which they investigated the value of the Bernreuter Personality Inventory for the direct prediction of total scholastic achievement on the secondary school level. They conclude, "The data at hand furnish no evidence that the Bernreuter inventory is measuring any traits that contribute in any important degree to successful achievement in the high school."[3]

Despite the fact that certain variables may be of little value for purposes of direct prediction, they may be more valuable as prognostic of differential ability. This is primarily due to the fact that in finding a differential correlation coefficient a negative and a positive direct relationship coefficient may be brought together and the effect is an additive one. A survey of Tables I and II reveals that of the 24 coefficients of correlation between the Bernreuter Personality Inventory scores and the honor point averages, fourteen are positive and ten are negative; however, there are only six opportunities for negative

[2] F. H. Finch and C. L. Nemzek, "The Relationship of the Bernreuter Personality Inventory to Scholastic Achievement and Intelligence," *School and Society*, 36: 594–596, 1932.

[3] *Ibid.*, 596.

and positive direct relationship coefficients to have an additive effect to produce higher differential coefficients of correlation.

Despite the fact that the direct relationship coefficients are so low that they almost preclude any significant differential relationship coefficients, the value of the Bernreuter scores for purposes of differential prediction was determined by Segel's[4] method. In Tables IV and V are included the differential prediction coefficients based upon the data available for the 92 boys and the 99 girls under consideration. In Table VI are presented the means and standard deviations in step interval units as they were substituted in Segel's formula for obtaining differential prediction coefficients.[5]

A study of the data based upon the boys indicates that the Bernreuter Personality Inventory has no value for predicting the differences between the honor point averages for mathematics and English (5–6). The differences correlate $-.066 \pm .070$, $.054 \pm .070$, and $-.054 \pm .070$ with B1–N scores, B2–S scores, and B4–D scores, respectively. Not any of these coeffi-

[4] David Segel, Differential Diagnosis of Ability in School Children. Baltimore: Warwick and York, Inc., 1934. Pp. 86. David Segel, Prediction of Success in College. Office of Education Bulletin 1934, No. 15. Washington, D. C.: Government Printing Office, 1934. Pp. 98.

David Segel, "The Construction and Interpretation of Differential Ability Patterns," *Journal of Experimental Education*, 2: 283–287, 1934.

David Segel, "Differential Prediction of Ability as Represented by College Subject Groups," *Journal of Educational Research*, 25: 14–26, January, 1932; 25: 93–98, February, 1932.

David Segel, "Differential Prediction of Scholastic Success," *School and Society*, 39: 91–96, January 20, 1934.

David Segel and J. R. Gerberich, "Differential College Achievement Predicted by the American Council Psychological Examination," *Journal of Applied Psychology*, 17: 637–645, December, 1933.

J. Murray Lee and David Segel, "The Utilization of Data from Simple or Direct Prediction in the Development of Regression Equations for Differential Prediction," *Journal of Educational Psychology*, 24: 550–554, October, 1933.

[5] The formula is:

$$r_{(a-b)x} = \frac{r_{ax}\sigma_a - r_{bx}\sigma_b}{\sqrt{\sigma_a^2 + \sigma_b^2 - 2r_{ab}\sigma_a\sigma_b}}$$

TABLE IV

*Differential Prediction Coefficients of Correlation Based upon (1) Intelligence Quotients, (2) B1–N Scores, (3) B2–S Scores, (4) B4–D Scores, (5) Honor Point Averages in Mathematics, (6) Honor Point Averages in English, (7) Honor Point Averages in History and Social Science, and (8) Honor Point Averages in Languages for 92 Boys**

VARIABLES	1	2	3	4
5–6	.017	−.066	.054	−.054
5–7	−.156	−.101	.186	−.042
5–8	−.087	−.124	.174	.111
6–7	−.329	−.069	.253	.181
6–8	−.109	−.073	.144	.184
7–8	.069	−.037	.005	.088

* The probable errors of these differential prediction coefficients range from .063 to .070. These probable errors were calculated by the usual formula. It has been suggested by Dr. George K. Bennett, of the Psychological Corporation, that the probable errors thus derived are too small. He recommends the use of the probable error of the difference between r's as a more accurate method of obtaining the probable errors. The larger probable errors would demonstrate more decisively the lack of differential predictive value. The same conclusion applies to the data in Table V.

TABLE V

*Differential Prediction Coefficients of Correlation Based upon (1) Intelligence Quotients, (2) B1–N Scores, (3) B2–S Scores, (4) B4–D Scores, (5) Honor Point Averages in Mathematics, (6) Honor Point Averages in English, (7) Honor Point Averages in History and Social Science, and (8) Honor Point Averages in Languages of 99 Girls.**

VARIABLES	1	2	3	4
5–6	.086	−.088	.026	−.054
5–7	.073	−.024	−.070	−.111
5–8	−.098	−.098	.026	.130
6–7	.010	.061	−.130	−.161
6–8	−.184	−.041	.010	.150
7–8	−.161	−.073	.089	.226

*The probable errors of these differential prediction coefficients range from .064 to .068.

TABLE VI

Means and Standard Deviations in Step Interval Units of (1) Intelligence Quotients, (2) B1–N Scores, (3) B2–S Scores, (4) B4–D Scores, (5) Honor Point Averages in Mathematics, (6) Honor Point Averages in English, (7) Honor Point Averages in History and Social Science, and (8) Honor Point Averages in Languages for 92 Boys and 99 Girls.

VARIABLES	BOYS (N = 92)		GIRLS (N = 99)	
	Mean	S.D.	Mean	S.D.
1	9.83	2.36	9.49	2.44
2	18.32	2.86	18.95	3.32
3	22.58	1.83	21.19	2.08
4	23.35	2.44	22.85	2.36
5	5.26	1.39	5.11	1.47
6	5.11	1.28	5.82	1.30
7	5.59	1.47	5.45	1.59
8	5.21	1.57	5.80	1.71

cients are as large as their probable errors; hence, they do not have statistical significance.

The differences between the honor point averages for mathematics and history and social science (5–7) correlate $-.101 \pm .070$, $.186 \pm .068$, and $-.042 \pm .070$ with B1–N scores, B2–S scores, and B4–D scores, respectively. It is apparent that these coefficients are not statistically significant. When intelligence is held constant by means of partial correlation, the relationship between the B2–S scores and the differences between the honor point averages for mathematics and history and social science changes from $.186 \pm .068$ to $.193 \pm .068$. This partial coefficient is 2.84 times its probable error; however, it has practically no value for prediction purposes.

The differences between the honor point averages for mathematics and the languages (5–8) correlate $-.124 \pm .069$, $.174 \pm .068$, and $.111 \pm .069$ with B1–N scores, B2–S scores, and B4–D scores, respectively. These coefficients indicate that the Bernreuter Personality Inventory scores are of negligible value for predicting differences in achievement between mathematics and languages.

The differences in achievement between English and history and social science (6–7) correlate $-.069 \pm .070$, $.253 \pm .066$, and $.181 \pm .068$ with B1–N scores, B2–S scores, and B4–D scores, respectively. When intelligence is held constant by means of partial correlation, the relationship of $.253 \pm .066$ between B2–S scores and differences in achievement between English and history and social science becomes $.280 \pm .065$. This coefficient is slightly more than four times its probable error, but it is very doubtful if it has any decided practical value. The differential prediction coefficient $r_{1(6\text{-}7)}$, is $-.329 \pm .063$; whereas, the direct prediction values, r_{16} and r_{17}, are $.460 \pm .055$ and $.538 \pm .050$, respectively. These data certainly demonstrate that the intelligence quotients are more effective for direct prediction than for differential; furthermore, the Bernreuter Personality Inventory scores have negligible value for direct or differential prediction.

The differences in achievement between English and the languages (6–8) correlate $-.073 \pm .070$, $.144 \pm .069$, and $.184 \pm .068$ with B1–N scores, B2–S scores, and B4–D scores, respectively. No statistical significance may be attached to these coefficients. The same conclusion applies to the relationships obtained when the differences in achievement between history and social science and languages (7–8) are correlated with B1–N scores, B2–S scores, and B4–D scores, the coefficients being $-.037 \pm .070$, $.005 \pm .070$, and $.088 \pm .070$, respectively.

The data presented in Table IV have failed to reveal any value that the Bernreuter Personality Inventory may have for purposes of differential prediction. This finding is also corroborated by the results based upon the girls as shown in Table V. In conclusion one may say that the present investigation indicates that the Bernreuter Personality Inventory has no practical value for purposes of direct or differential prediction of academic achievement under the conditions existing among students at the University High School of the University of Minnesota.

SUMMARY AND CONCLUSIONS

The purpose of this study was to reveal the value of the Bernreuter Personality Inventory for purposes of direct and differential prediction. Data were available for 92 boys and 99 girls, and the following eight variables were used:

(1) intelligence quotients,
(2) B1–N scores,
(3) B2–S scores,
(4) B4–D scores,
(5) honor point averages in mathematics,
(6) honor point averages in English,
(7) honor point averages in history and social science, and
(8) honor point averages in languages.

These data demonstrate that the Bernreuter Personality Inventory has negligible value for direct or differential prediction.

USE OF THE INTERNATIONAL SCORING MACHINE FOR THE RAPID COMPUTATION OF TABLES OF INTERCORRELATIONS

G. FREDERIC KUDER
The University of Chicago

ROYER and Toops[1] have developed a rapid method for obtaining correlations from Hollerith cards on which the geometric codes of scores for each variable have been punched according to a method described by Dunn.[2] The Royer-Toops method calls for punching a card for each individual, and requires the use of a Hollerith key-punch and a counter-sorter. By the simple expedient of preparing an answer sheet and stencil for each *code*, it is possible to use the International Scoring Machine for obtaining the correlations. This substitution is not recommended if Hollerith equipment is easily available. The technique will, however, be found to be much more efficient than the usual manual methods if intercorrelations of at least five variables are to be obtained. Its use is most efficient when the number of cases is between fifty and two hundred.

The detailed outline of the method follows.

1. Prepare a list of the subjects and their raw scores on the variables involved. Allow two blank columns after the list of subjects, and five narrow blank columns after the raw scores for each variable.

[1] Royer, E. B., and Toops, H. A. "The Statistics of Geometrically Coded Scores," *Journal of the American Statistical Association*, Vol. 28, 1933, pp. 192–198.

[2] Dunn, Halbert L. "Adaptation of New Geometrical Code to Multiple Punching in Mechanical Tabulation," *Journal of the American Statistical Association*, Vol. 27, 1932, pp. 279–286.

In Table 1, the subjects have been listed in column (1), and the raw scores of variables X, Y, and Z have been listed in columns (4), (10), and (16), respectively.

2. Divide the range of the raw scores of each variable into not more than sixteen class intervals. Assign grouped scores of from 0 to 15 to these class intervals. Prepare tables or formulas for transmuting raw scores into grouped scores.

The formula developed for transmuting Variable X into grouped scores is $X = \frac{X'}{3} - 2$ in which X = grouped score, and X′ = raw score. Fractions are dropped in using this formula. The same result is produced by developing a table. In this case the table is:

Raw scores	*Grouped score*
51–53	15
48–50	14
45–47	13
42–44	12
39–41	11
36–38	10
33–35	9
30–32	8
27–29	7
24–26	6
21–23	5
18–20	4
15–17	3
12–14	2
9–11	1
6– 8	0

The formulas used for grouped scores Y and Z are as follows:

$$Y = \frac{Y'}{10} - 1 \qquad Z = \frac{Z'}{5} - 3$$

Fractions are always dropped in using these formulas. Different formulas or tables must, of course, be developed for each variable.

3. Enter in the second column for each variable the grouped score corresponding to the raw score.

These have been entered in columns (5), (11), and (17) in Table 1.

4. Label the next four columns for each variable 1, 2, 4, and 8, respectively. Then for each grouped score, place checks in the following columns so that the sum of the values of the checks exactly equals the grouped score. Not more than one check should be placed in a column.

TABLE 1

Coding of Scores

(Adapted from Royer and Toops)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
Individual	Position on answer sheet		Variable X						Variable Y							Variable Z				
	Question Number	Answer Position	Raw Score	Grouped Score	Score codes				Raw Score	Grouped Score	Score codes					Raw Score	Grouped Score	Score codes		
					1	2	4	8			1	2	4	8				1	2	4	8
A	1	1	40	11	✓	✓	.	✓	98	8	.	.	.	✓		78	12	.	.	✓	✓
B	2	1	32	8	.	.	.	✓	115	10	.	✓	.	✓		49	6	.	✓	✓	.
C	3	1	18	4	.	.	✓	.	100	9	✓		.	✓		19	0	.	.	.	.
D	4	1	50	14	.	✓	✓	✓	72	6	.	✓	✓	.		80	13	✓	.	✓	✓
E	5	1	7	0	.	.	.	.	87	7	✓	✓	✓	.		40	5	✓	.	✓	.
F	6	1	38	7	✓	✓	✓	.	59	4	.	.	✓	.		64	9	✓	.	.	✓
G	7	1	33	9	✓	.	.	✓	46	3	✓	✓	.	.		71	11	✓	✓	.	✓
H	8	1	51	15	✓	✓	✓	✓	86	7	✓	✓	✓	.		73	11	✓	✓	.	✓
I	9	1	10	1	✓	.	.	.	31	2	.	✓	.	.		36	4	.	.	✓	.
J	10	1	36	6	.	✓	✓	.	150	14	.	✓	✓	✓		33	3	✓	✓	.	.
K	11	1	42	12	.	.	✓	✓	108	9	✓	.	.	✓		72	11	✓	✓	.	✓
L	12	1	30	8	.	.	.	✓	160	15	✓	✓	✓	✓		64	9	✓	.	.	✓
M	13	1	45	13	✓	.	✓	✓	62	5	✓	.	✓	.		90	15	✓	✓	✓	✓
N	14	1	38	10	.	✓	.	✓	147	13	✓	.	✓	✓		54	7	✓	✓	✓	.
O	15	1	35	9	✓	.	.	✓	91	6	.	.	.	✓		67	10	.	✓	.	✓
P	16	1	24	6	.	✓	✓	.	24	1	✓	.	.	.		28	2	.	✓	.	.
Q	17	1	13	3	.	✓	.	.	139	11	✓	✓	.	✓		23	1	✓	.	.	.
R	18	1	16	3	✓	✓	.	.	83	7	✓	✓	✓	.		64	9	✓	.	.	✓
S	19	1	21	5	✓	.	✓	.	15	0	.	.	.	.		55	8	.	.	.	✓
T	20	1	27	7	✓	✓	✓	.	136	12	.	.	✓	✓		46	5	.	✓	✓	.

Columns (6) to (9) have been labeled 1, 2, 4, and 8, respectively, for Variable X. Columns (12) to (15) and (18) to (21) have been labeled in the same way for Variables Y and Z.

The X-score for individual A is 11; it is therefore represented by checks in score code columns 1, 2, and 8, since the sum of these numbers is 11. A's Y-score is 8; this score is therefore represented by a check in code column 8 for the Y-scores. The score of 12 in Variable Z is represented by checks in code columns 4 and 8. All grouped scores in the table have been coded in this fashion.

5. For this and following steps, International Scoring Machine answer sheets are used (although the actual entries made, as outlined below, are codes rather than answers).

See Figure 1 for an illustration of part of an answer sheet. This represents the left portion of the standard form for 150 questions with five alternatives each. The large figures are question numbers; the small figures at the five spaces after each question number indicate the answer positions.

Assign each individual an answer position on the answer sheet. Write the question number in column (2) and the answer position in column (3). It is recommended that positions be assigned in sequence by columns on the answer sheet.

In Table 1, the first subject, A, has been assigned position 1–1 (answer position 1 after question number 1 on the answer sheet). Subject B has been assigned position 2–1, and Subject T has been assigned position 20–1. If there were more subjects, the 31st subject would be given position 1–2 at the head of the second column of answer positions, and the 60th subject would be given position 30–2, at the foot of the second column. The 61st subject would be given position 1–3 at the head of the third column, and so on.

6. Label a separate answer sheet for each of the code columns of each variable.

In the case of Table 1, 12 answer sheets are prepared, labeled, respectively, X–1, X–2, X–4, X–8, Y–1, Y–2, Y–4, Y–8, Z–1, Z–2, Z–4, and Z–8.

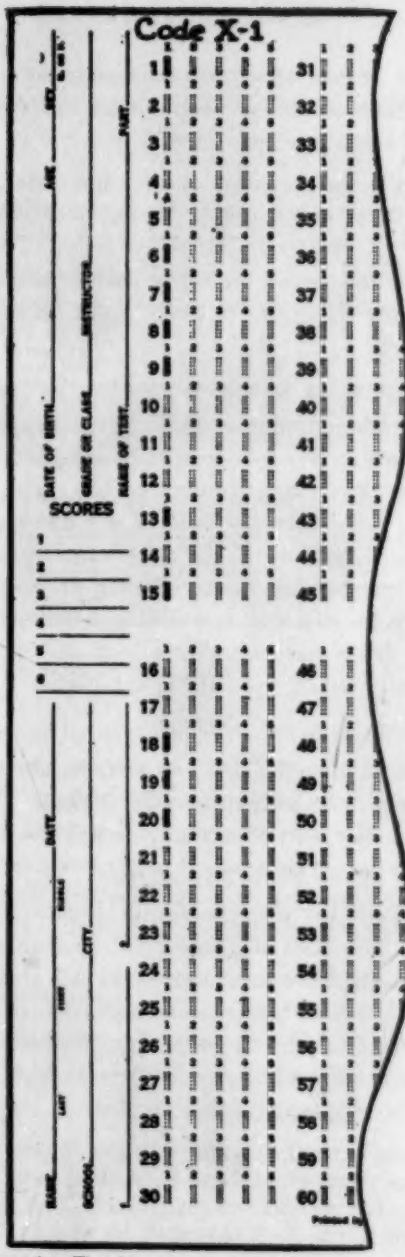


FIGURE 1. Entries on answer sheet for code X–1.

7. Fill in the answer sheet for each code column by blackening the positions of those people who are checked in the column. Keep these answer sheets in order.

For example, the "answer sheet" for code X–1 as illustrated in Figure 1 has the following positions blackened: 1–1, 6–1, 7–1, 8–1, 9–1, 13–1, 15–1, 18–1, 19–1, and 20–1.

8. Make up code stencils for every answer sheet, each stencil having a hole corresponding to every mark on the corresponding answer sheet.

This means making twelve stencils for the example used.

9. Prepare a table similar to Table 2A, listing the variables in order and allowing five spaces in each direction for each variable. Also allow two extra columns at the right. In labeling the rows at the left, leave the first of the five spaces for each variable unlabeled, and write the code numbers of the answer sheets for each variable in the remaining four spaces in order of size from top to bottom. In labeling across the tops of the columns, write these same numbers, from left to right, in the first four spaces, leaving the fifth space blank.

This has been done in Table 2A.

10. Put the first stencil (X–1) in the machine and score all the answer sheets in order (for number "right"). The scores are entered, in order, in the column for code X–1 of the table.

See column (2) in Table 2A.

Then score all the sheets with the second stencil (X–2), entering the scores in column X–2 of Table 2A. Continue in this manner until all sheets have been scored by all stencils, and the table is complete. Note that each combination is entered twice in the table, except for the diagonal entries which are the number of positions blackened on each answer sheet. The diagonal entries should be obtained twice, as a check.

In the example used, the entry in the upper left cell is the number of positions which have been blackened on sheet X–1. The entry in the second cell in column X–1 is the number obtained when sheet X–2 is scored by stencil X–1.

TABLE 2A

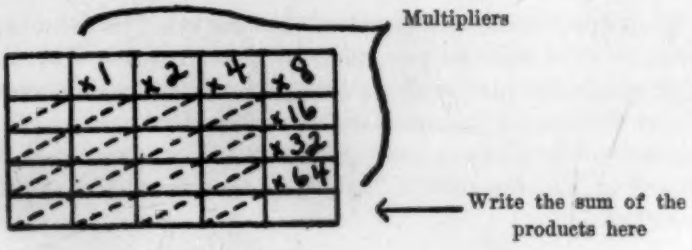
Table of Numbers of Entries Common to Each Pair of Code-Columns in Table 1

(Adapted from Royer and Toops)

(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Code		Variable X					Variable Y					Variable Z					Total on sheet	Σ
		1	2	4	8		1	2	4	8		1	2	4	8			
X	1 . .	10	5	5	5	___	4	4	5	3	___	5	5	4	8	___	10	
	2 . .	5	10	6	4	___	5	5	7	5	___	7	5	4	5	___	10	
	4 . .	5	6	10	4	___	5	3	6	4	___	6	6	3	6	___	10	
	8 . .	5	4	4	10	___	6	5	5	6	___	7	7	5	8	___	10	
Y	1 . .	4	5	5	8	___	11	6	6	5	___	9	6	3	6	___	11	
	2 . .	4	5	3	5	___	6	10	6	4	___	8	4	4	5	___	10	
	4 . .	5	7	6	5	___	6	6	10	4	___	9	5	5	6	___	10	
	8 . .	3	5	4	6	___	5	4	4	10	___	5	6	4	4	___	10	
Z	1 . .	5	7	6	7	___	9	8	9	5	___	12	6	4	8	___	12	
	2 . .	5	5	6	7	___	6	4	5	6	___	6	10	4	5	___	10	
	4 . .	4	4	3	5	___	3	4	5	4	___	4	4	8	3	___	8	
	8 . .	8	5	6	8	___	6	5	6	4	___	8	5	3	11	___	11	

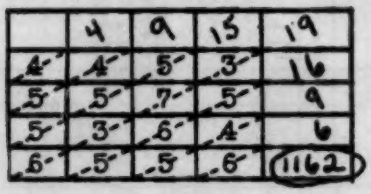
11. Check the table for symmetry about the diagonal running from upper left to lower right. Column X–1 reading from top to bottom should have the same entries as row X–1, reading from left to right, and so on.

12. In each block of 16 cells obtain the sum of the entries along each diagonal running from lower left to upper right; multiply each diagonal total by the amount given in the following diagram, and obtain the sum of these products. Enter this figure in the cell at the lower right corner of the block, and draw a ring around it.



The total obtained is ΣXY, the sum of the products of the grouped scores of the variables listed at the end of the row and the head of the column. In the case of the diagonal blocks this total is ΣX^2, the sum of the squares of the one variable involved.

Take as an example the group of sixteen cells in the X-row and Y-column of Table 2A. These are added along the diagonals, and the sums listed as indicated:



The operations involved are

Multiplier		*Sum on diagonal*		
1	×	4	=	4
2	×	9	=	18
4	×	15	=	60
8	×	19	=	152
16	×	16	=	256
32	×	9	=	288
64	×	6	=	384
		Total	=	1162

These steps can be performed in sequence on the calculating machine and the total obtained without clearing. If the multipliers are memorized, this operation proceeds rapidly. Step 12 has been performed for all variables in Table 2B.

13. Copy the diagonal entries of the table in a "total" column at the right. Multiply each entry by the code value of the row and obtain the sum of these four products for each variable. Enter this sum in the next column, and circle the entry. This operation should be performed twice. The quantity so obtained is ΣX, the sum of the grouped scores of the variable involved.

TABLE 2B

Use of Table 2A for Obtaining Data for the Computation of Intercorrelations

(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Code		Variable X					Variable Y					Variable Z					Total on sheet	Σ
		1	2	4	8		1	2	4	8		1	2	4	8			
			10	10	30	32		4	9	15	19		5	13	16	35		
X	1 . .	10	5	5	5	15	4	4	5	3	16	5	5	4	8	15	10	
	2 . .	5	10	6	4	5	5	5	7	5	9	7	5	4	5	11	10	(150)
	4 . .	5	6	10	4	10	5	3	6	4	6	6	6	3	6	5	10	
	8 . .	5	4	4	10	(1470)	6	5	5	6	(1162)	7	7	5	8	(1394)	10	
			4	9	15	19		11	12	32	32		9	14	16	30		
Y	1 . .	4	5	5	6	16	11	6	6	5	16	9	6	3	6	16	11	
	2 . .	4	5	3	5	9	6	10	6	4	9	8	4	4	5	10	10	(151)
	4 . .	5	7	6	8	6	6	6	10	4	10	9	5	5	6	9	10	
	8 . .	3	5	4	6	(1452)	5	4	4	10	(1485)	5	6	4	4	(1093)	10	
			5	13	16	35		9	14	16	30		13	12	13	24		
Z	1 . .	5	7	6	7	15	9	8	9	5	16	12	6	4	8	15	12	
	2 . .	5	5	6	7	11	6	4	5	6	10	6	10	4	5	6	10	(152)
	4 . .	4	4	3	5	5	3	4	5	4	6	4	4	8	3	11	8	
	8 . .	8	5	6	8	(1395)	6	5	6	4	(1093)	8	5	3	11	(1489)	11	

In Table 2B, the diagonal entries (corresponding to the number of entries on each answer sheet) have been entered in column (17). The sum of the four products of these entries and the code values of their rows have been entered in column (18) for each variable.

14. Proceed with the calculation of the intercorrelations, making use of the circled entries.

An efficient systematic procedure for reducing the quantities obtained in the second table to an intercorrelation table has been described by Toops.[3]

It should be noted that the amount of labor involved in the preparation of stencils and scoring can be reduced by at least a third if the computer is willing to use eight class intervals instead of sixteen. Only three code sheets and stencils are then necessary for each variable.

[3] Toops, Herbert A. "Computing Intercorrelations of Tests on the Adding Machine," *Journal of Applied Psychology*, Vol. 6, 1922, pp. 177 ff.

If a large number of runs is called for, a set of answer sheets should be made up for each thirty runs. It has been general experience that if an answer sheet is put through the machine much oftener than this, inaccuracies are likely to occur. The extra sets of answer sheets are quickly made, however, by using the stencils as guides in filling out the new sheets.

AN ANALYSIS OF EIGHT MEASURES OF PERSONALITY AND ADJUSTMENT IN RELATION TO RELATIVE SCHOLASTIC ACHIEVEMENT*

NEAL E. DROUGHT
University of Chicago

THE rapid increase in college enrollment figures during the post-war years with the resulting increase in student mortality gave rise to many attempts to provide an index of probable scholastic achievement. (For an excellent summary of the more noteworthy of these studies see *1*: 1–40). The majority of these studies have used either a test of scholastic aptitude or a measure of previous achievement (such as high school average or percentile rank in class) or both. The correlations between predicted and achieved grades have quite regularly fallen within the range of plus .55 and plus .70. The fundamental assumption of the present study is that the discrepancies between predicted and achieved grades implied by these correlations are to be accounted for in some measure by social and emotional factors.

With a view toward obtaining a better understanding of these factors, two measures, each with four subdivisions, were administered to 750 freshman men entering the College of Letters and Science of the University of Wisconsin in September, 1935. The two scales employed were the *Adjustment Inventory* by Hugh Bell[1] and the *Wisconsin Scale of Personality*

* The study of which this represents a part was made possible by a grant from the research committee of the Graduate School of the University of Wisconsin and received aid in the form of clerical help from the Works Progress Administration. The writer is indebted to Professor Kimball Young for continued assistance and encouragement in this study.

[1] *The Adjustment Inventory* by Hugh M. Bell. Published by the Stanford University Press, Stanford University, California. 1934.

Traits by Ross Stagner.[2] Those acquainted with these measures will recall that the Bell Inventory provides measures of Home, Health, Social and Emotional adjustment. The four divisions of the Stagner scale are described by the author as follows:

> "On scale W a high score indicates a person who is nervous, moody, tends to worry over trifles, and who is likely to adopt defense mechanisms. (Nervousness and emotionality.)
>
> "On scale X a high score indicates an individual who has a highly subjective orientation, a tendency to daydream, and a susceptibility to criticism. (Introversion.)
>
> "On scale Y a high score indicates high self-esteem and self-confidence.
>
> "On scale Z a high score indicates an individual who is averse to social contact, prefers intellectual to social or athletic activities, and takes a passive rôle. (Social passivity or seclusiveness.)"[3]

The criterion of scholastic achievement used in this study was the discrepancy between predicted and achieved grade point averages. At the University of Wisconsin the predicted average is obtained by a regression equation having as variables the student's percentile rank in his high school class and his percentile rank on the American Council Test of Scholastic Aptitude administered during freshman orientation week. (See 2.)

The achieved average is determined by assigning three grade points per credit for a grade of "A", two grade points per credit for a "B", one per credit for a "C", zero for a "D", minus one-half for an "E", and minus one for a failure, and then dividing total credits into total grade points.

Thus by using the discrepancy between predicted and actual grades as a criterion of achievement rather than the attainment of an arbitrary standard such as a 1.00 ("C") average we have in some measure controlled the factor of scholastic

[2] *The Wisconsin Scale of Personality Traits* by Ross Stagner. Published for the author by the Bernat Publishing Co., Chicago. 1935.

[3] From a communication from the author.

aptitude. In terms of this definition a student with a grade point prediction of 2.00 or "B" who achieves a 1.00 or "C" average is just as much a "failure" as is a student having a prediction of 1.00 ("C") who attains an average of .00 ("D"). Similarly it is possible to pick out those students whose achievement is abnormally high in relation to their prediction.

PROCEDURE

The first step was to determine the correlation between grade point discrepancies and scores on each of the subdivisions of the Bell and Stagner scales.

The following list shows that in all cases these correlations were of low order.

	Correlation
First semester grade point discrepancy and home adjustment (Bell A)	+.03
health adjustment (Bell B)	−.05
social adjustment (Bell C)	+.05
emotional adjustment (Bell D)	−.01
Bell total	+.01
nervousness (Stagner W)	+.05
introversion (Stagner X)	+.07
self-esteem (Stagner Y)	−.01
social passivity (Stagner Z)	+.15

In no case was there a suggestion of reliable relationship. This is in essential agreement with previous studies employing similar measures.

The second step in the analysis was to determine to what extent the scales differentiated between students with discrepancies of at least plus .60 grade points (1½ PE) and those with at least minus .60 grade points. These distributions were obtained from an unselected sample of 250 cases from the original group of 750. The scores for each of the 90 students in the plus group and the 115 in the minus group were recorded for each of the eight personality measures and the mean for each of these distributions was determined.

TABLE I

Distribution of Scores on the Four Subdivisions of the Bell Inventory for Students Whose Discrepancies Were Plus or Minus .60 Grade Points

SCORE	HOME ADJUSTMENT		HEALTH ADJUSTMENT		SOCIAL ADJUSTMENT		EMOTIONAL ADJUSTMENT	
	Plus .60	Minus .60	Plus .60	Minus .60	Plus .60	Minus .60	Plus .60	Minus .60
28 and over	0	0	1	1	4	0	3	2
26–27	1	1	1	0	0	3	1	0
24–25	1	1	0	0	3	2	2	1
22–23	2	0	0	0	4	4	3	4
20–21	1	0	1	1	6	7	1	1
18–19	3	3	0	1	5	4	1	1
16–17	3	4	1	0	8	3	8	11
14–15	5	2	3	4	6	10	5	7
12–13	5	5	6	5	7	11	4	9
10–11	3	7	7	6	13	16	6	10
8– 9	5	10	14	17	11	12	13	18
6– 7	14	12	16	23	7	14	14	9
4– 5	17	21	19	28	5	18	9	18
2– 3	13	24	13	23	8	11	10	14
0– 1	17	25	8	6	3	0	10	10
Total	90	115	90	115	90	115	90	115
Mean	7.1	5.8	7.0	6.4	12.4	10.8	9.5	9.1

For example, the average score on the home adjustment scale for students in the plus group was 7.1 and for the minus group 5.8. The difference between these means is 1.3 with a probable error of .58. In a similar fashion the same data were secured on each of the remaining seven scales. In no case did the difference between the means equal or exceed four times the probable error of the difference.

Similarly, as a third step, the average grade point discrepancies were determined for students with ratings of "excellent" and "unsatisfactory" on each division of the Bell Inventory and for students scoring in the lowest and highest deciles on the four divisions of the Stagner scale. Naturally the number of cases varied from scale to scale.

In this analysis it was found that the difference between

TABLE II

Distribution of Scores on the Four Subdivisions of the Stagner Scale for Students Whose Discrepancies Were Plus or Minus .60 Grade Points

SCORE IN PERCENTILE	NERVOUSNESS		INTROVERSION		SELF-ESTEEM		SOCIAL PASSIVITY	
	Plus .60	Minus .60	Plus .60	Minus .60	Plus .60	Minus .60	Plus .60	Minus .60
91–100	19	10	16	9	15	17	19	12
81– 90	7	6	13	9	6	5	5	6
71– 80	9	12	6	10	6	13	11	3
61– 70	1	11	4	9	10	9	6	15
51– 60	8	2	8	7	8	14	11	12
41– 50	9	17	11	14	9	7	9	18
31– 40	7	14	10	12	11	20	8	20
21– 30	9	11	8	17	8	8	0	0
11– 20	7	11	5	11	9	8	6	9
0– 10	14	21	9	17	8	14	15	20
Total	90	115	90	115	90	115	90	115
Mean	51.9	44.4	55.4	44.6	52.2	50.7	54.9	46.2

the mean grade point discrepancies was at least four times the probable error of this difference for students scoring high and for those scoring low on the X and Z scales of the Stagner measure indicating introversion and social passivity respectively. On scale X the difference between the means was $.29 \pm .07$, and on scale Z the difference between the means was $.30 \pm .06$.

Apparently then introversion and social passivity are in some measure related to scholastic achievement. An examination of the distributions, however, indicates that we should be confusing the issue by simply resorting to these statistical devices without further analysis.

The final step in this section of the study was, then, to analyze the tests item by item. This was done by determining the differences in the percentage of "yes", "?", and "no" responses on each part of the Bell and Stagner scales given by students who achieved above prediction and those who fell

TABLE III

Distribution of Grade Point Discrepancies of Students Scoring High and Low on the Four Subdivisions of the Bell Inventory

GRADE POINT DISCREPANCY FOR FIRST SEMESTER	HOME ADJUSTMENT		HEALTH ADJUSTMENT		SOCIAL ADJUSTMENT		EMOTIONAL ADJUSTMENT	
	0–1	11 and over	0–2	11 and over	0–4	20 and over	0–2	16 and over
1.20 and over	3	2	3	2	0	1	0	3
1.00– 1.19	1	3	2	0	1	2	2	2
.80– .99	10	5	5	5	6	3	5	7
.60– .79	6	13	2	8	6	11	3	7
.40– .59	9	7	4	5	7	10	4	7
.20– .39	7	5	6	7	7	8	6	4
.00– .19	14	11	10	10	12	9	8	6
– .01– – .19	14	10	8	10	4	9	10	5
– .20– – .39	9	8	9	9	18	7	9	12
– .40– – .59	7	6	7	11	13	14	11	15
– .60– – .79	10	5	5	7	6	5	8	10
– .80– – .99	5	5	6	4	5	7	4	6
–1.00– –1.19	5	2	4	2	3	2	5	1
–1.20– –1.39	0	1	0	0	2	1	2	0
–1.40– –1.59	2	0	1	1	1	0	0	0
–1.60– –1.79	3	1	2	0	1	1	1	3
Total	105	84	74	81	92	90	78	87
Mean	–.06	.07	–.11	–.04	–.14	–.03	–.19	–.10

below prediction. Selecting from the Bell Inventory the ten items showing the most marked differentiation, we should describe the *poor* student as one who does not often feel self-conscious because of his personal appearance, is in doubt as to whether or not he is often the center of favorable attention at a party, has not frequently had to keep quiet or leave the house in order to have peace at home, does enjoy social dancing a great deal, does not get discouraged easily, does like to participate in festival gatherings and make "whoopee," does hesitate to enter a room by himself when a group of people are sitting around the room talking together, does not sometimes envy the happiness others seem to enjoy, does not cross the

TABLE IV

Distribution of Grade Point Discrepancies of Students Scoring High and Low on the Four Subdivisions of the Wisconsin Scale of Personality Traits

GRADE POINT DISCREPANCY FOR FIRST SEMESTER	PERCENTILE RANK ON SCALE OF							
	NERVOUSNESS		INTROVERSION		SELF-ESTEEM		SOCIAL PASSIVITY	
	0–10	91–100	0–10	91–100	0–10	91–100	0–10	91–100
1.20 and over	2	2	0	0	0	1	1	1
1.00– 1.19	1	3	1	3	1	1	0	2
.80– .99	7	3	6	4	0	5	7	5
.60– .79	4	12	2	9	6	8	5	15
.40– .59	5	4	5	6	5	5	4	7
.20– .39	7	3	5	2	2	4	7	7
.00– .19	11	11	10	5	9	3	6	13
– .01– – .19	6	2	9	7	2	3	9	8
– .20– – .39	13	8	12	4	14	12	13	7
– .40– – .59	10	12	13	5	6	6	10	13
– .60– – .79	7	4	4	3	5	5	8	4
– .80– – .99	3	6	6	5	5	3	7	3
–1.00– –1.19	6	0	5	1	1	5	3	1
–1.20– –1.39	1	1	1	0	1	1	0	0
–1.40– –1.59	1	1	0	0	1	2	1	1
–1.60 and under	3	1	1	0	1	1	4	0
Total	87	73	80	54	59	65	85	87
Mean	–.15	.00	–.19	+.10	–.20	–.12	–.21	+.09

street to avoid meeting somebody, and does take the responsibility for introducing people at a party.

It will be noted that six of the items involve social participation, three relate to emotional adjustment and one purports to refer to home adjustment.

With reference to the Stagner scale the ten items showing the greatest differentiation would suggest that the *poor* student does not like to be by himself a great deal, does not indulge much in daydreams, does not like to have quiet amusements rather than exciting ones, does not get tired of work

quickly, does not like to spend his vacations at some quiet place rather than at a lively resort, does not think much about it when praised, is not often considerably upset when unsuccessful in any undertaking, does not usually think a great deal before deciding anything, would not like to listen to a lecture on classical music, and does not usually talk spontaneously. Five of these reflect social passivity, four introversion, and one emotionality or nervousness.

On the basis of these ten selected items from each measure, the scales were scored for a group of 563 students for whom the necessary data were available. Each of the items was considered a unit value since no weighting scheme seemed warranted at this stage of development. A high score was associated with poor achievement.

The scores for the two sets of ten items were combined to give a single index and the correlation between these selected-items scores and grade point discrepancies was + .20.

CONCLUSIONS

It becomes apparent that the tests used in this study are of little value for the purpose of differentiating between students who will obtain grades above prediction and those who will fall markedly below prediction.

In regard to the Wisconsin Scale of Personality Traits this conclusion is simply a verification by a new technique of the results of previous investigators using similar scales. Stagner himself has made one of the most extensive investigations of the relationship between intelligence, achievement, and personality factors and found all correlations to be low (*3*). The present data do not, however, corroborate his conclusion that "Extreme personality measures seem to counterbalance advantages in aptitude, making for equal achievement in opposed groups" (*3:* 10). Were this true, we should expect a significant negative correlation between percentile rank on emotionality or self-sufficiency and grade point discrepancy. Such was not the case.

Insofar as extreme scores on the parts of the Wisconsin Scale may be taken to indicate adjustment or maladjustment the same criticism may be applied as is apparent in regard to the Adjustment Inventory. The present data show quite clearly that there is no relationship between adjustment as measured and relative performance in college. And yet there is abundant literature to support the thesis that social and emotional adjustment, for example, do play an important part in college achievement. It is unquestionably true that maladjustment may show itself by abnormally high achievement as well as by failure, but an examination of the grade point discrepancies for students with high and low scores on the Adjustment Inventory shows no bimodal distribution such as would be expected were the scales consistent in pointing out maladjustment.

It thus becomes clear that a different approach will be necessary if the problem of obtaining an index of personal-social-emotional factors in college achievement is to be solved.

SUMMARY

In an effort to account for the personal-social-emotional factors operative in scholastic achievement, the Adjustment Inventory by Hugh Bell and the Wisconsin Scale of Personality Traits by Ross Stagner were administered to 750 freshman men entering the College of Letters and Science of the University of Wisconsin in September, 1935. The scores on the four subdivisions of each of these measures were correlated with the discrepancy between predicted and achieved grades. These correlations were uniformly low.

There was no significant difference in the average scores on any of the eight scales between students whose achieved grades were markedly below prediction and those whose grades were markedly above prediction.

In comparing the average grade point discrepancies of students with high and with low scores on each of the eight parts of the two measures it was found that there was a significant

difference in average achievement between students who scored high and low on introversion and on social passivity as measured by the Stagner scale. The students scoring high on introversion and social passivity had the higher average grades.

The two measures were analyzed item by item and there were selected the ten items from each showing the greatest difference between students who worked above prediction and those who worked below prediction. When a group of 563 tests were scored on the basis of these items, the correlation between these scores and grade point discrepancies was found to be +.20.

These measures are, therefore, of limited value for the purpose of this study. It is apparent that a new approach will have to be worked out for obtaining a measure of those personal-social-emotional factors related to academic achievement.

REFERENCES

1. BERGSTRESSER, JOHN L., *The Prediction of Academic Achievement for Men in the College of Letters and Science*, Ph.D. thesis, University of Wisconsin, 1936.
2. HENMON, V. A. C., AND HOLT, F. O., *A Report on the Administration of Scholastic Aptitude Tests to 34,000 High School Seniors in Wisconsin in 1929 and 1930*, Bulletin of the University of Wisconsin, 1936.
3. STAGNER, ROSS, "The relation of personality to academic aptitude and achievement," *Journal of Educational Research*, 26: 648–60, 1933.

MENTAL ABILITIES AT SENESCENCE: A SURVEY OF PRESENT-DAY RESEARCH

GEORGE LAWTON
Columbia University

OF the various life epochs that have been studied by scientific investigators, old age has received the least attention. In recent years, however, the problems of the aged have begun to force themselves upon us because of socio-economic factors. A declining birth-rate and an improvement in medical skill have combined to raise markedly the average age of the population since the turn of the century. But while the life-span has been lengthened, the work-span has been curtailed and though large sums are being spent in the attempt to support the improvident aged, their numbers are constantly being increased by economic conditions.

Nor should the problem be limited to keeping the aged alive, even assuming that our provisions for their economic and physical welfare were adequate. If we are to discover and provide what old people need in order to be most useful and happy, it is important that we make accurate determinations of their mental abilities and emotional lives, and discard many of the traditional and folk-lore judgments now current. But while the problems of the aged are manifold: economic, medical, emotional, psychiatric, educational and occupational, there is very little integrated research to guide us in a consideration of these problems. It is the purpose of this paper to sketch the highlights of psychometric contributions to the science of geriatrics now in its birth-throes.

Most of the early treatments of senescence were concerned with physiological and philosophical techniques for postponing decline and death. Typical of this approach is Mechnikov's

"The Prolongation of Life" (1908) and Lorand's "Old Age Defined, The Cause of Old Age and Its Postponement by Hygienic and Therapeutic Measures" (1912). Since men have always shown a greater interest in life's duration than in its quality, such books appear with unfailing regularity. Serge Voronoff's "Life, A Study of the Means of Restoring Vital Energy and Prolonging Life" and M. Gallichan's "Youthful Old Age; How to Keep Young" are recent variations on the same theme.

One of the first psychometric studies was Beeson's "Intelligence at Senescence" (1920). Twenty cases, of average age 75, in a highly selected home for the aged, yielded an average IQ of 83 on the Stanford-Binet. This low score Beeson attributed to decline in mental processes at senescence. No deterioration, however, was found with advancing age, as the older half of the men proved more intelligent than the younger half. The best performance of the group was in vocabulary, a fact which has since played a very important part in the study of senescent mental efficiency. However, the size and select character of Beeson's group alone reduce the value of his results. Foster and Taylor in "The Applicability of Mental Tests to Persons over 50 Years of Age" (1920) reported the use of the Yerkes-Bridges Point Scale to test a number of younger and older normal and psychotic persons. The study showed among the normal individuals a general falling off with age. But its chief importance lies in making a suggestion of the greatest methodological value for future research: namely, that senescents of the same level of intelligence should be compared if we are going to measure mental efficiency, and that the determination of the intellectual status of an older person would be more accurate if two mental ages were given: one comparing him with his normal contemporaries and the other comparing him with his own adolescent ability (or that of normal young people).

G. Stanley Hall's "Senescence" in 1922 is the most elaborate attempt by a professional psychologist to view decline in its totality, but though it is valuable as a move to take the subject

out of the realm of popular stereotypes, especially with respect to the emotional lives of the aged, it is nevertheless an armchair performance based on the literature of senectitude and on returns to a questionnaire. As for psychometric studies, Hall makes no mention of them, dismissing the subject with these words: "There are no mental tests of general recognized validity above the 'teens,' so that we have no criteria for determining psychological age for even the elderly" (p. 196).

The Stanford University Studies in Later Maturity made under the direction of W. R. Miles (now of the Institute of Human Relations at Yale) are the most extensive investigations that have been undertaken thus far of the relationship of certain human abilities to age. A number of papers published since 1931 report comparisons made on a total of 2500 subjects between the ages of 7 and 94 (250 of these over 70). The intelligence measure used was the Otis Higher given as a fifteen minute speed test. The maximum scoring ability, whether the test be of general intelligence or of a specific activity, occurs between the ages of 18 and 49, usually in the earlier half of this period; 10 to 25 per cent of those in later maturity and old age do as well in most tests as the average in middle maturity. Motor abilities were found to reach their maximum in early adulthood (between 18 and 29). In the results for 400 subjects with the speed factor eliminated, Miles showed that speed did not slacken before mental power. In a later paper, however, C. C. Miles reported that comparisons of the curtailed Otis, speeded and unspeeded indicated that in early and middle maturity speed declined faster than power, whereas in late maturity the decline of power was relatively more apparent. In general, it was felt that there was a diminution in reaction speed and sum of energy available for new work types rather than in intelligence, decline correlating with physiological rather than with psychological changes. It is important to observe that feelings of insecurity and inferiority are considered the greatest handicap; appropriate training and practice,

persistence and other stable personality factors are considered assets to mental longevity.

The chief criticism of the Stanford Studies is that the sampling was a select one, the subjects having been secured through church groups, clubs, lodges, and philanthropic societies, and that they were divided according to educational levels, which cannot be taken as a real separation in terms of native ability. The Studies show definitely the gradual decline of large groups of the population in general ability and in specific activities, the variability of maturation and age-decline in separate abilities and even the variability of individuals on the different tests, but they cannot tell how much an individual has dropped below the efficiency to be expected from one of his intellectual level. Because 10 to 25 per cent of older people score as well as, or better than, the average of those in their prime of efficiency does not necessarily mean that they have managed to maintain this efficiency better than their contemporaries; they may, in fact, have had much greater capacity to start with and have lost far more than many of those in the same age group who obtained the same or worse scores. On the other hand, a man of sixty testing very low in comparison with twenty year olds might by nature be of inferior intellect but of preserved efficiency.

The chief problem, then, in measuring deterioration at senescence is to compare the mental efficiency of various old people and still manage to control the original intellectual level so that the comparison is between persons of equal endowment. But we cannot use educational levels or occupational backgrounds as ways of insuring intellectually comparable groups. We must in some manner obtain an estimate of the senescent's peak performance whenever it occurred. While the most definitive method would be a follow-up of mental power scores of individuals from shortly after birth through senescence, this is not yet a feasible procedure, and it is necessary to devise a reliable alternative.

The earliest to offer such an alternative technique was H. Babcock (An Experiment in the Measurement of Mental Deterioration, 1930; Dementia Praecox, A Psychological Study, 1933). Her "Test of Mental Efficiency" designed for use with psychotic patients is based on the assumption that deterioration occurs first in new learning and last in earliest formed material. Terman's vocabulary is used as a measure of intellectual level and a series of speed and learning tests is employed for the efficiency phase of intelligence. The discrepancy between the two types of scores yields a measure of the deterioration.

But, it may be asked, what justification has Babcock to assume that the vocabulary remains constant throughout the life span? It is true that other writers have noted that language is relatively slow to deteriorate (Miles), that it has a high correlation with mental age (Terman), and that an extensive vocabulary does not seem to be dependent on education (O'Connor). All of this shows that vocabulary is an important index of intelligence. But the language ability of the aged may be too good to serve as an accurate control. What if it improves with advancing age? O'Connor pointed out that although vocabulary does not alter materially after maturity, there *is* an increase of two school years in the period 23–50. If this is so and we attempt to measure the discrepancy between vocabulary score and the average score of a series of other tests in order to determine failing efficiency, a false deterioration will be shown unless we apply a correction. But how much this should be and how it is to be arrived at are very controversial points. Proponents of the constancy thesis call attention to the difficulty of believing that the aged begin to fail in memory, in learning and in motor skills and improve only in a single ability: vocabulary.

Jeanne Gilbert tackles this crucial aspect of the problem of deterioration in "Mental Efficiency at Senescence" (1935) using a revision of the Babcock Test made by the latter. Gilbert selected two groups, an experimental one consisting of 175

individuals of average age 63.8 and a control group of 185, average age 25. All subjects had clear sensory functions, spoke English and had no history of psychotic episodes. Illiterates were excluded. The groups were comparable in occupational status and in motivation (at least so Gilbert assumed). The younger and older groups were matched for vocabulary level. Then 33 short tests requiring timed recall of old association, quick perception, various types of memory, motor ability new learning, etc., were administered. The scores of the older persons on the efficiency tests were compared with those of the younger individuals who had received the same vocabulary score (*i.e.*, were of the same intellectual level). The discrepancy of the two average scores, or "efficiency index," was taken as that presumably between the aged person's ability at senescence and as it was when he was younger. This "index" is of unusual importance in the study of the aged individual. To know the average score on the efficiency tests means little until we discover how the individual's performance compares with that which is expected from one of his intellectual level. He might be of natively inferior intelligence but functioning with good efficiency or endowed with high intellect but functioning with very poor efficiency.

Though Gilbert believes that her own study confirms the claim that vocabulary remains constant after maturity has been reached, for the sake of safety she offers her efficiency indices in two forms: in one table there is no correction for a possible improvement in vocabulary; in the other, going O'Connor one better, she marks down in vocabulary all of the older groups by three years instead of two, though she says such a large correction is unnecessary.

However, in both tables, there is shown a large and reliable difference in the mental efficiency of persons in the sixties and persons in the twenties. The efficiency index of no young person falls as low as the mean of the older group and only two older persons reach the mean of the younger group.

There is a definite decrease in efficiency for all levels of intellect but there is a marked tendency for those of the very highest levels of intellect (largest vocabulary) to show relatively less decrease of efficiency than those of less intellect. Apparently the greater the intellect with which one is originally endowed, the less he tends to deteriorate or the longer he tends to retain his original equipment of efficiency.

One important result of Gilbert's study is to show that individuals vary in efficiency and that age alone is an insufficient factor by which to judge one's ability to meet the requirement of a particular vocation. This study is in line with the results of other similar investigations: an arbitrary retirement age of 65 ought to be supplanted by a more flexible system. While some sixty year olds are unable to handle their tasks with their usual ability, others can pursue their work with accustomed efficiency, still others can work, but at a somewhat lower level. But even if the older worker has lost a small measure of the speed which was his at an earlier age, this may be compensated for by additional experience, judgment, etc. If an injustice is done by our industrial system by arbitrarily excluding persons in the sixties from labor which some might still do, what shall we say of the compulsory elimination from industry of the forty-year old worker which shelves persons still in their prime?

Gilbert's study is an important one and it is therefore worth time to consider its shortcomings. Apart from the fact that much more work must be done before the constancy of vocabulary can be accepted as proved and that the results may be subject to revision with regard to the correlation of mental age and vocabulary in the fifties and sixties, to the applicability of the Babcock test to normal subjects, to the selection of the battery for determining efficiency, the chief criticism is that the older and younger group are not strictly comparable. Gilbert states that 76.5 per cent of the older group was unemployed at the time of the study, whereas only 38.9 per cent of the younger group were so situated, but dismisses as a problem for

future research "The relationship between unemployment and inefficiency, particularly among older persons in whom deterioration is expected."

However, certain pertinent questions must be raised. What effect has unemployment and the consequent lack of proper work habits and practice on the ability to handle the thirty-three tests of efficiency? What about the dissimilarity of the younger and older groups in motivation and emotional reaction to the tests? Employment or unemployment, regardless of age, are significant items to be matched if we want comparable groups. Moreover, we have considerable justification for believing unemployment or a previous history of long or frequent unemployment periods is an even more serious threat to the morale of older persons than to that of younger ones. In this respect, Gilbert's study suffers from a limitation which is true of psychometric work generally in the field of old age: little consideration is given to the extent to which emotional sets and personality patterns may affect test results. Perhaps it is too much to ask that groups be equated with respect to freedom from serious neurotic involvements as well as from a history of psychotic episodes. It is not too much to ask for specific information on how each young and old person reacted to the test situation as a whole and to the individual tests. Such qualitative information is of peculiar importance in the study of old people.

The writer is of the opinion that subsequent research will show that economic and social insecurity play a large part in hastening mental deterioration in the forties and fifties and are particularly responsible for the marked increase in the deterioration rate of the late sixties over the early sixties, which Gilbert herself notes but leaves for the future to explain. If psychometric studies at any age level are of limited usefulness when socio-economic and personality factors are ignored, at senescence such isolated test results are of dubious validity.

Adult Intelligence, A Psychological Study of Test Performances by Weisenberg, Roe and McBride, published in 1936,

reports the results of a battery of language intelligence, educational achievement and non-language tests given to a total of 70 men and women ranging in age from 20 to 59, mean age 36.1. Because of the age span, this volume does not belong, strictly considered, in a survey of the research done in the field of senescence. However, the approach to the problem and the testing program, for the most part, are appropriate to older subjects and for that reason the study ought to be cited here.

One very important methodological novelty of the study is the technique used to secure an unselected sample of the population. Theoretically, the best method would be to test, in house-to-house fashion, every individual in a given community. This is not generally feasible, though Jones and Conrad in "The Growth and Decline of Intelligence, A Study of a Homogeneous Group between the Ages 10 and 60" (1933), made a survey of almost a complete New England community. Instead, the authors of "Adult Intelligence" worked in the surgical and orthopedic wards of three Philadelphia hospitals, making certain, of course, that the physical disabilities were not of a type to affect test performances. The hospital group besides furnishing a random sample of average and near-average persons, employed as well as unemployed, also provided subjects who were grateful for devices to help pass the time away and who were in nowise limited as to the amount of time that they could devote to examinations or interviews. The problem of securing a group of suitable subjects is especially acute at the senescent level, and the idea of a hospital group is a very useful one.

Unlike most other studies of adult intelligence, the results of which are based on one or two short tests, or at any rate a relatively small number, the battery employed here was a very extensive one, consisting of 34 tests and involving 15–20 hours of testing time. Furthermore, the hospital situation permitted a more friendly and personal relationship and a more intimate knowledge of qualitative factors.

The conclusions of "Adult Intelligence" are in accord with that of most other studies in the field. The authors find that the greatest extent of mental development is reached before the twenties and from this decade through the fifties there is little further gain and comparatively little decline. We again are shown that the peak of development and the course of abilities with age are different for different performances or abilities, and that there is considerable variation among individuals in achievement. Sex differences are negligible. Formal schooling is not as important a factor in determining test results as is generally supposed. Interestingly enough, in working out the correlations of various tests with age, the only evidence of an increase in score with age appears in the vocabulary tests, notably the Thorndike Word Knowledge. (This is confirmation of the Gilbert position referred to earlier relative to the correction which it is necessary to make for the improvement in vocabulary in later maturity.) The most marked evidence of decrease in scores with age occurs in the analogies and the non-language tests.

Dr. Irving Lorge in "The Influence of the Test upon the Nature of Mental Decline as a Function of Age" (1936), makes a valuable contribution to the methodology of studying senescence. In order to determine the relationship between age and various tests of mental ability, eleven tests were administered to a group of 143 adults ranging in age from 20 to over 70. Some of these were tests of power, others a mixture of power and speed. Lorge maintains that the existence of a high correlation, in some cases as high as .85 between power and power-speed tests of ability (*i.e.*, between IER Intelligence Scale CAVD and Otis Self-Administering Test of Mental Ability Higher Examination, 20 minute time allowance), is not sufficient to prove that each test measures the same thing. Indeed, the relationship of each type of test to age differs markedly.

The results of Lorge show that the relationship between age and ability varies as the test used to measure the ability. By matching similar scores on CAVD in three age groups (20–25,

27.5–37.5, over 40) with scores obtained on the other tests, he was able to determine the penalty which age imposed upon the particular measure of ability which is employed, in this case a mixed power-speed test. He applies his findings to the Jones and Conrad study already cited in this article and to the Miles and Miles investigation, "The Correlation of Intelligence Scores and Chronological Age from Early to Late Maturity" (1932), employing in the case of the former a correction of .780 points per year after age 20 and in the later study a correction of .480 points. In both instances, he finds that the adjusted means do not show the same gradual decline exhibited in the obtained scores.

Lorge concludes that speed obscures mental power per se in older adults and a correction must be applied to take care of those physiological, educational or emotional changes that interfere with the older person's power to cope with mental tasks. Such a correction may alter the curve of mental decline to a curve of mental plateau or even to a curve of mental growth, though Lorge admits that mental growth in the later age brackets may be more the result of the special selection of death (death coming earlier to the physically and mentally mediocre) than a true growth. The reported facts of mental decline as concomitants of age are more apparent than real and are due to the fact that the test used to measure mental ability is not a genuine measure of mental power.

We have seen that the problem of evaluating senescence or deterioration cannot be met without the use of tests that have been standardized on adults and for which adequate norms, at different age levels, are available. Dr. David Wechsler's "Bellevue Intelligence Examination," to be published in the fall of 1938, will make available for the first time data of this kind, having norms for all ages up to 70. Any discussion of this test, however, must be postponed until its official publication.

In an address made in October, 1937, before the New York Neurological Society, Dr. Wechsler offered certain observations

concerning the approach to the study of senescence which are of considerable importance. He pointed out that "normal deterioration" is hardly different qualitatively or quantitatively from the deterioration regarded as pathological by psychiatrists and neurologists. For deterioration, however defined, does not begin in senescence or senility but starts immediately after the point of maximal virility. This point for most human abilities and most individuals begins somewhere in the twenties. Mental deterioration of the "normal" variety, *i.e.*, not caused by disease processes or injury, does not appear suddenly as a precipitate drop in the developmental curve. It is a gradual process, though as aging progresses, it may increase in momentum. This point is well brought out by Wechsler in Chapter 7 of his book *The Range of Human Capacities* (1935). It is worth noting that Dr. Alfred E. Cohn of the Rockefeller Institute working in the field of physical aging, principally heart disease, makes exactly the same point concerning the gradual decline in the functioning of the heart ("The Aging of the Heart Muscle Regarded from a General Biologic Point of View," 1929).

We may sum up our discussion of the mental functioning of old people by saying no adequate study of their mental abilities and performances yet exists. The first requisite for such a study is a satisfactory psychometric tool; a testing program which will compare the old person not only with his normal contemporaries but with his own abilities as they were when he was young. In this testing program due consideration must be given to two facts: first, different test performances reach their peak at different ages and show different courses through maturity; second, even those performances which seem to decline rapidly after maturity in relation to other test performances are in reality seen to decline slowly when compared to the rate of development. Hence our norms must be based on large numbers of persons of every age level from early to late maturity and senescence.

Each type of activity must be measured by itself at sufficient length and with a sufficient number of cases properly selected to afford a reliable estimate. We still have to learn something of the relationship of various mental abilities at the adult level, to say nothing of the senescent level with all its special problems.

But even when an adequate psychometric approach will have been devised, it will have to be properly applied: *i.e.*, the part that emotional, social and other non-intellectual factors play in affecting test results must be suitably evaluated. Test results simply as test results are even more suspect at the senescent level than at earlier periods.

Facts about the mental abilities of old people are of great importance in mental hygiene and in social planning. Only when we have reliable data here, can we make a just decision as to the work which old people can do, and as to which ones can continue working past the technical retirement age, which ones should enter upon a modified work program at such a time and finally which persons should be given simpler tasks even before the retirement age has been reached. Our knowledge today about the possibilities of such modified work programs for older people is very scanty and unreliable. In a planned society no individual is completely scrapped. Depending on the nature and extent of the deterioration every individual can perform some socially useful work no matter how elementary or limited it may be.

JUDGING EXPRESSIVE MOVEMENT: III. JUDGMENTS OF DOMINANCE-FEELING FROM PHONOGRAPH RECORDS OF VOICE[1]

PHILIP EISENBERG AND ELEANOR ZALOWITZ
Columbia University

THE PROBLEM

WE have previously studied the question how accurately and in what ways people judge dominance-feeling from the expressive movements, handwriting and gait.[2] In this paper we are continuing the investigation of the general problem of judging personality from expressive movement by studying judgments of dominance-feeling from phonograph records of voice, another experimentally isolated expressive movement. This investigation is of special importance, since, in forming our impressions of our fellows from their expressive movements, voice plays a large role. Here we are not interested in the content of speech, but in the manner and qualities of speech, for example, loud or soft voice, fast or slow, smooth or hesitant, even or uneven tempo, etc. Though this may seem to imply that we analyzed isolated features of voice, actually we permitted our judges to form an impression from the entire voice.

[1] The authors are indebted to Dr. Otto Klineberg and Dr. Max Hertzman for their advice and assistance.

[2] Eisenberg, P. Judging Expressive Movement: I. Judgments of Sex and Dominance-Feeling from Handwriting Samples of Dominant and Non-Dominant Men and Women. *Journal of Applied Psychology*, 1938, Vol. 22, No. 5, pp. 480–486. Eisenberg, P, and Reichline, B. B. Judging Expressive Movement: II. Judgments of Dominance-Feeling from Motion Pictures of Gait. *Journal of Social Psychology* in press.

As a secondary problem, which we have not investigated previously, we studied the consistency of judgment over a short period of time.

PROCEDURE

On the basis of the Maslow "Social Personality Inventory," a rating scale of dominance-feeling (both of these being administered twice) and an interview, 8 extremely dominant and 8 extremely non-dominant Barnard College women were selected from a group of 238.[3] We obtained a one minute record of each subject reading a description of a character from Mann's "The Magic Mountain." The subjects were well aware that records were being made of their voices, but instead of trying to conceal the fact, we preferred to utilize this impressive situation believing that it would even more serve to differentiate the extremes.

The records were played for a group of judges, 43 psychology students at Brooklyn College,[4] in the following random order: D, ND, D, D, ND, ND, D, D, ND, D, ND, D, ND, D, ND, ND. After the playing of each record, the judges were asked to decide whether the girl was dominant or non-dominant. After an interval of 30 minutes, during which judgments were made from motion pictures,[5] the records were played in another random order, and judgments were made. At the end of each presentation, the judges were asked to write what were considered to be the characteristics of voice of dominant and non-dominant women. The following definition of dominance-feeling was given to the judges:

[3] For an elaboration of the method of selecting extremes of the distribution of dominance-feeling, see Eisenberg, P. Expressive Movements Related to Feeling of Dominance. *Arch. Psychol.*, 1937, No. 211.

[4] We are grateful to Dr. Austin B. Wood for his assistance in securing subjects.

[5] Subjects were required to judge dominance-feeling from motion pictures of facial expression, but due to a serious error in splicing the films, these data could not be used.

"Feeling of dominance is defined as general self-confidence, and high level of self-evaluation. It includes a forcefulness of character, 'strong' personality, a feeling of being 'on top,' a feeling of mastery over others, a feeling of general capability. It may assume several forms in different people: despotism, aggressiveness, helpful leadership, guidance, or just feeling dominant with no active expression. . . . Low dominance-feeling, in general, consists of the reverse or absence of these feelings. It includes feelings of general inferiority, lack of self-confidence, uncertainty, feeling dominated by others, timidity and shyness, often a willingness to be self-effacing, and self-consciousness."[6]

RESULTS

1. How well can people judge dominance-feeling from phonograph records of voice? Referring to Table I, we see that in

TABLE I

Average Judgments of Dominance-Feeling for Two Presentations

PRESENTATION	N	AVERAGE ERRORS	PER CENT CORRECT	C*
I	43	5.93	62.94	.26
II	40	6.25	60.94	.22

* Vernon's method of computing coefficients of contingency was used, *op. cit.*

both presentations dominance-feeling was judged a little better than average chance. However, we must consider this only a tentative conclusion until we make a further analysis. (See question 5.) It is interesting to compare our C's of .26 and .22 for both presentations with a C of .29 which Allport and Cantril[7] obtained when voices were matched with scores on the Allport A-S test. However, we cannot make a direct comparison

[6] For an elaboration of the concept of dominance-feeling, see Maslow, A. H. Dominance-Feeling, Behavior and Status. *Psychol. Rev.*, 1937, 44, 404–429.

[7] Allport, G. W. and Cantril, H. Judging Personality from the Voice. *J. Soc. Psychol.*, 1934, 5, 37–55. See also, Vernon, P. E. The Matching Method Applied to Investigations of Personality. *Psychol. Bull.*, 1936, 33, 149–177.

since Allport and Cantril utilized different experimental conditions and another personality trait.

2. How consistent are the judgments of voice over a period of time? Referring to Table I, we see that there were slightly more errors made in judging the second time. However, the difference between the average number of errors of both presentations is not statistically reliable (critical ratio .97). The slight increase in errors was most probably due to less attention because of fatigue and boredom.

However, the correlation between both presentations for total number of errors was only $.42 \pm .09$. As a reliability figure this is exceedingly low and we might reasonably conclude from this that the judgments of voice were rather inconsistent.

But because of the peculiar conditions of this experiment, it is necessary to analyze consistency in absolute rather than relative terms, that is, in terms of actual identity of response for each judge and each record instead of a correlation measure. When we make this analysis, we find that 69.43 per cent of all responses were identical in both presentations. From this we may more reasonably conclude that judgments of voice are fairly consistent, that is, that the judgments are not made by chance but on the basis of some definite criteria. Just what those criteria are we shall see later (question 5). It is very doubtful that memory played any significant part in this identity of judgment.

3. Do women judge better than men? Comparing 13 men with 30 women in the first presentation, it was found that there was no statistically reliable difference in the average number of errors made between these two groups (critical ratio .89). That women can judge dominance-feeling from phonograph records of voice no better than men must remain a tentative conclusion in the light of the small number of cases. However we obtained the same result in connection with judgment of gait, and other experiments in judging support this conclusion.[8]

4. In what respects do the good and poor judges differ? The judges were required to write free descriptions of what they

[8] Eisenberg and Reichline, *op. cit.* Vernon, *op. cit.*

considered to be the characteristic differences between the voices of dominant and non-dominant women. In making a qualitative comparison of the extremely good and extremely poor judges, we arrive at the following tentative conclusion: The good judges tend to give a fuller, more detailed description than the poor judges. Their descriptions are more objective and directed to the subtler aspects of expression. However, these differences are very slight and must be considered with caution, due to the small number of cases.

5. *In judging individual records, do the judges reveal stereotyped attitudes?* We have answered only partially the question of how well individuals judge dominance-feeling from phonograph records of voice, and in what ways they arrive at their judgments. An analysis of individual records may throw some further light on these questions.

When we correlate the judgments of individual records for both presentations we obtain a correlation of .75 $\pm$ .08, indicating that though individuals may differ to some extent when making judgments from time to time, the group as a whole is making about the same judgments for individual records. This suggests that there is something about the record and the stereotyped attitude the judges bring with them that influences their judgment to uniformity.

Table II indicates that there is a wide range in the ease of judging any record. We have combined all the judgments for each record for both presentations and computed the reliability of the difference from chance, whether the judgments were correct or incorrect. (This combination is justified by the correlation of .75 above.) It was found that records 1, 2, 4, 10, 13, 14, 15 and 16 were judged reliably better than chance; records 3, 5 and 11 reliably worse than chance; and the other five records just about chance, using a critical ratio of 2.50 as a dividing line.

When we examine the records themselves, we find that the judges are influenced by the records themselves and their preconceptions rather than by the actual personalities of the women

TABLE II

Reliability of Judgment of Individual Records

RECORD	NO. ERRORS	PER CENT CORRECT*	D FROM 50%	σD	D/σD
1	24	71.08	21.08	7.41	2.84
2	2	97.59	47.59	5.74	8.29
3	60	27.71	22.29	7.37	3.00
4	23	72.29	22.29	7.37	3.00
5	70	15.66	24.34	6.78	3.59
6	39	53.01	3.01	7.75	.39
7	30	62.86	12.86	7.63	1.69
8	56	32.53	17.47	7.52	2.32
9	46	44.58	5.42	7.73	.70
10	8	90.36	40.36	6.37	6.33
11	68	18.07	31.93	6.93	4.61
12	27	67.47	17.47	7.52	2.32
13	2	97.59	47.59	5.74	8.29
14	24	71.08	21.08	7.41	2.84
15	16	80.72	30.72	6.99	4.39
16	3	96.39	46.39	5.86	7.91

* Of 83 possible judgments.

judged. We find in the eight records judged better than chance that four are dominants and four non-dominants. The four dominants happen to have voices which fit the preconception of dominant voice, that is, their speech is dramatic, well-modulated, non-hesitant, in general, "good" speech, and also possessing mature qualities of voice, not shrill and high-pitched. The non-dominants happen to have the opposite qualities of voice. Of the three judged reliably worse than chance, No. 3, a dominant, has all the stereotype characteristics of a non-dominant, and No. 5 and 11, non-dominants, the characteristics of dominant voice. Those judged just about chance have some dominant and some non-dominant characteristics so that the subject will be called a dominant or non-dominant depending on what the judges feel is most important at the moment of judging. For example, No. 7, a dominant, has all the characteristics of good speech, but her voice is immature, even childish and

shrill; whereas No. 12, also a dominant, has a mature and deep voice, but she does not read well, since she hesitates and falters frequently.

It seems then that the judges did not really know what was the characteristic expression of a dominant girl as distinguished from a non-dominant girl, and were judging by the most obvious characteristics, expressing their stereotyped attitudes largely on whether the speech was "good" or "bad" speech, and on the maturity of the voice, these being synonymous for them with dominant and non-dominant speech. It is also probable that in enumerating the characteristics of the dominant as opposed to the non-dominant, they were expressing this stereotyped attitude rather than the subtleties of expression. This tends to confirm our suspicion that the good judges were those who were more attentive to these subtle aspects of expression, in spite of the fact that they could not verbalize any better than the poor judges.

This analysis brings us to the conclusion that we obtained judgments a little better than chance merely because of a few records "easy" to judge. And we find in Table III that when

TABLE III

Percentage of Incorrect Judgments After Eliminating the Records "Easiest" to Judge

NO. OF RECORDS ELIMINATED	PER CENT INCORRECT	
	Presentation I	Presentation II
0	37.21	39.06
1	39.68	41.67
2	42.36	44.46
3	45.26	47.69
4	48.26	51.04

we eliminate one by one the four "easiest" records to judge, we approach chance judgments very rapidly, considering 50 per cent as chance. It might also be added, that if the judged groups were matched for reading ability and maturity of voice, the judges would have made more nearly chance judgments.

6. What is the relationship between social agreement and correctness of judgment? When we correlate the number of errors with the number of identical judgments made for individual judges we obtain a correlation of $-.09 \pm .11$ for the first presentation and $.38 \pm .09$ for the second. In the light of the change of sign and the size of the P.E.'s we may consider this a zero relationship.

However, when we correlate the percentage of correct judgments made on individual records with the percentage of identical judgments combined for both presentations we obtain a correlation of $.53 \pm .12$. The percentage of identical judgments for individual records can be considered a measurement of social agreement since the group as a whole is making about the same judgments and revealing that they are basing their judgments on the same preconceptions. If that is so, we may conclude from the above correlation that social agreement is to some extent a measure of accuracy of judgment. But when we examine Table IV we find that the records judged reliably better

TABLE IV

Relation Between Reliability of Judgment of Records and Identity of Judgment Over a Period of Time

RELIABILITY OF JUDGMENT	NO. OF RECORDS	AVERAGE % IDENTITY
Better than chance	8	78.71
Chance	5	52.97
Worse than chance	3	72.07

than chance and the records judged reliably worse than chance have the highest percentage of identical judgments, whereas the records judged about chance have the lowest percentage of identical judgments. If there were an equal number of records judged better than chance and worse than chance, then the above correlation would probably have been zero.

These results agree with the findings of Allport and Cantril,[9]

[9] Allport and Cantril, *op. cit.*

Taylor[10] and Stagner[11] who had subjects judge various aspects of personality from voice. They conclude that the degree of social agreement is greater than the degree of correctness of judgment, and that this is due to the fact that judgments are based on a preconceived stereotyped attitude. Taylor states: "There is a tendency for auditors to be most consistent (with each other) in their judgments when they agree least with the subject's self-ratings. It is at least clear that their degree of consistency in judging a given trait bears no positive relationship to their accuracy in judging that trait."

DISCUSSION

We have found that people judge dominance-feeling from phonograph records of voice, just as in the case of judgment of handwriting and gait, a little better than chance, but when we analyze individual records we find that they make their judgments on the basis of preconceived notions which may or may not correspond to reality. It is not necessary to repeat here our discussion of judgment of gait, but all our remarks are relevant to the judgment of voice.

It is interesting, in summary, to compare the judgments of handwriting, gait and voice. All judgments were made of extremely dominant and non-dominant women, and though different groups of judges were used the following percentages of correct judgments were obtained: handwriting 64.00 per cent, gait 63.89 per cent, and voice 62.94 per cent. The similarity in percentages is rather striking. If the similarity is not coincidental, it can be interpreted to mean (1) that the judges, using the same preconception of dominance, arrive at approximately the same decisions regardless of the expressive movement judged, and (2) that there is some element of reality in this stereotype but not enough for prediction or to be able to

[10] Taylor, H. S. Social Agreement on Personality Traits as Judged from Speech. *J. Soc. Psychol.*, 1934, 5, 244–248.

[11] Stagner, R. Judgments of Voice and Personality. *J. Ed. Psychol.*, 1936, 27, 272–277.

use the matching method as a technique for measuring the interrelationship between various aspects of personality.

These interpretations seem plausible, since we, as well as other investigators, have found a high degree of social agreement in judgment of expressive movement. That social agreement bears no relation to correctness of judgment is an important finding and should disturb the faith of those who believe that if we have enough judges who agree we can learn the truth. This is nothing more than another illustration of Sherif's[12] thesis that any group of people living under the same cultural conditions will develop social norms, not only of conduct, but of perception and judgment, which will tend to make these psychological processes uniform, but not necessarily realistic.

This conclusion becomes even more plausible when we find what this preconception or social norm is. In a previous study[13] it was found that almost all subjects, both men and women, consider dominance desirable and non-dominance undesirable. With this as a major clue, we find that our material illustrates a principle of judgment discovered by Asch, Block and Hertzman[14] in connection with judgments of photographs, professions, political figures and political slogans, namely that "the judgments of a single situation are related to each other by a person in accordance with an underlying attitude of acceptance or rejection." This is perhaps most clearly expressed in the judgment of voice where "good" speech and mature voice was considered dominant speech and "poor" speech and immature voice, non-dominant. In the judgment of gait, a smooth gait, with an even tempo, associated with an attractive-looking girl, obviously desirable characteristics, was considered to be dominant gait. Palmer-method handwriting was considered to be non-dominant handwriting, but Palmer writing is definitely

[12] Sherif, M. The Psychology of Social Norms. Harper, 1936.

[13] Eisenberg, P. *Arch. Psychol., op. cit.*

[14] Asch, S. E., Block, H. and Hertzman, M. Studies in the Principles of Judgments and Attitudes: I. Two Basic Principles of Judgment. *J. Psychol.*, 1938, 5, 219–251.

undesirable to our subjects, particularly if one recalls the tortures connected with it in elementary school. Desirable handwriting is distinguished, irregular and unusual, and that was considered to be dominant handwriting. In general, since there is a general underlying attitude of acceptance of dominance and rejection of non-dominance in the judgment of dominance-feeling from various expressive movements, all desirable qualities are ascribed to the dominant individual and undesirable qualities to the non-dominant individual.

In the light of all these data, one would hesitate to make judgments of personality from expressive movements alone. Yet the personnel manager or the employer often hires or rejects an applicant on just such flimsy evidence, by such impressions that teachers are found pleasing or displeasing by their students, etc. When we realize that these judgments are based on stereotype preconceptions and that the desires of the individual are projected into his judgment, we can well realize all the errors that are made. This does not mean that expressive movements cannot be used in describing the personality of an individual. Relationships have been found, and it may prove to be very useful in the study of behavior mechanisms, but more objective methods of measurement must be used and subtler analyses made before we can come to any conclusions. We cannot rely on the process of judgment alone, not even on the judgment of many people.

SUMMARY AND CONCLUSIONS

Phonograph records of the voices of 8 extremely dominant and 8 extremely non-dominant girls were played for a group of 43 college students to judge for dominance-feeling. After an interval of 30 minutes, the records were played again in a new order. It was found that:

1. Dominance-feeling was judged a little better than chance.

2. The judgments were found to be fairly consistent over a short period of time, for though the correlation between both

presentations for total number of errors was only .42, 69.43 per cent of the judgments were identical for both presentations.

3. Women did not judge any more accurately than men.

4. There was a slight tendency for good judges to make a more detailed description of voice, directed toward the subtler aspects of expression.

5. An analysis of individual records revealed that: (a) For both presentations, the group made approximately the same errors on individual records. (b) Wide variations in ease of judging records were found, 8 records being judged reliably better than chance, 5 about chance, and 3 reliably worse than chance. (c) When we analyze individual records we find that judgments were made chiefly on the basis of preconceived notions of expressive behavior, and that the good judges were more aware of the subtler aspects of expression. (d) If the three or four records "easiest" to judge are eliminated, we obtain judgments no better than chance.

6. There is no relation between social agreement and correctness of judgment.

7. A comparison of judgments on handwriting, gait and voice indicates that judgments are organized on the basis of a general attitude of acceptance of dominance and rejection of non-dominance, with a concomitant ascription of desirable qualities to the dominant individual and undesirable qualities to the non-dominant individual.

PREDICTION OF ACHIEVEMENT IN FOREIGN LANGUAGES

MAY V. SEAGOE
University of California at Los Angeles

PREDICTION of success in foreign languages is of major importance for two reasons. First there is the high proportion of failures that usually occurs with its waste of students' time, teachers' energy, and taxpayers' money. Second, more and more colleges are willing to waive the language requirement for students who have proved themselves in other fields. Language is no longer inevitable: it can be avoided by those for whom it has little value and who show little aptitude.

Many studies have been made in the field, most of them ably summarized by Kaulfers.[1] The median coefficients of correlation with bearing on the present study are .60 with prognosis tests, .49 with achievement in algebra, .46 with achievement in English, .385 with achievement in reading, .356 with intelligence tests, .164 with achievement in arithmetic, and −.24 with chronological age. The great range of coefficients obtained by various investigators, sometimes varying from low negative to nearly perfect positive correlation for a single characteristic, is noteworthy.

PROBLEM

The data available promised to shed some additional light on the subject, mainly because the method of administering tests more adequately represented the usual junior high school procedure than do most studies; guidance for ninth grade was based largely on test data gathered at entrance to the seventh grade. Some of the specific questions suggested were: (1) Is

[1] Kaulfers, Walter Vincent. "Present status of prognosis in foreign language." *School Review*, 1931, 39, 585–596.

the Luria-Orleans Prognostic Modern Languages Test superior to such general intelligence tests as Terman, Otis and Kuhlmann-Anderson in predicting success in languages? (2) Is the New Stanford Reading Test superior to a general intelligence test for this purpose? (3) Is it possible to establish fairly definite critical scores differentiating potential failure from satisfactory and satisfactory from recommended grades?

SUBJECTS, MATERIALS, AND PROCEDURE

The subjects were 120 students, 64 boys and 56 girls, whose records were being followed in connection with the Carnegie Guidance Experiment in Pasadena, California.[2] In this project all students entering 7–1 in September, 1934, were designated as the experimental group, a group which was followed for three years with a more than normally intensive guidance service.

The tests used in this study include the Terman Group Test given in May, 1934, just before seventh grade enrollment, the Kuhlmann-Anderson Intelligence Test given in May, 1936, and the remainder drawn from the general testing program carried on during the fall of 1934. Tests from the general program referred to in this article include the Otis Intermediate Examination, the New Stanford Achievement Test (sections on reading, physiology and arithmetic), the Luria-Orleans Modern Languages Prognosis Test, the Stenquist Mechanical Aptitude Test, and the Orleans Algebra Prognosis Test. The mathematics tests were included to determine their comparative relationship to or independence of the foreign language prediction, and the physiology and mechanical aptitude tests were included to explore the relation of foreign language achievement to such theoretically unrelated factors as scientific and practical ability.

Following the administration of each test the results were reported to the individual concerned and their significance for his educational and vocational plans was discussed. In case

[2] The Experiment was directed by Dean Grayson N. Kefauver of the School of Education at Stanford University.

the individual planned to take a language in the ninth year or was uncertain whether he should elect it, he was advised to enroll in the 8–2 Exploratory Languages course which gave six weeks each to Latin, Spanish, and French. At the end of the semester he made the decision concerning ninth grade language on the basis of his knowledge of his own abilities based on both experience and standardized test data, and on his educational plans. The choices were reviewed by the counselor on the same bases, and individuals whose choices seemed ill-advised were interviewed. The advice was based on the rule-of-thumb assumption that the result of the intelligence tests should indicate an IQ of 110 or higher, that the Luria-Orleans result should place the individual in the upper half of his class, and that the accomplishment in the exploratory languages course should be at least normal if the student is to succeed in a language. Such broad limits seemed justified in a school where approximately forty per cent of the pupils enter college. Exceptions were made where the pupil's home background, reputation as a student, apparent special aptitude in the field, or fixed vocational or educational objectives seemed to demand.

Teachers' marks at the completion of the first semester's work in the three languages constituted the criterion of success. They were stated only as "Recommended" or "Satisfactory" for the student's information in planning for tenth grade work since a personal letter took the place of the grade-report and the only official record was one of granting or refusing credit.

RESULTS

Derived rather than primary scores have been used because of their availability in intelligence and achievement tests. The result would theoretically be to underestimate relationships and critical ratios. The use of the intelligence quotient rather than mental age would be questionable under usual school conditions, but in this case a non-failure program has so reduced the differences in chronological age that the IQ is proportional to the mental age.

Intercorrelations between the tests analyzed have been presented in a recent article by the same author. The most significant fact is that the highest degree of relationship is among the intelligence tests, the next between intelligence and the verbal group, next between intelligence and the mathematics and physiology group, and that there is no relation between intelligence and mechanical aptitude. The r is $.63 \pm .03$ between reading achievement and language prognosis, $.73 \pm .03$ between Luria-Orleans and Terman, $.71 \pm .03$ between Luria-Orleans and Otis, and $.68 \pm .03$ between Luria-Orleans and Kuhlmann-Anderson. It is evident that there is an unreliable difference

TABLE I

Description of Groups

COMPARISON IN	SUB-GROUP	MEAN	SIGMA DIST.	SIGMA MEAN	RANGE
Chronological Age	Recommended	167.2	5.0	0.8	155–178
	Satisfactory	168.9	4.8	1.3	159–175
	Not Elected	172.4	6.7	0.8	159–197
Terman IQ	Recommended	122.4	13.3	2.3	100–149
	Satisfactory	111.1	6.7	1.8	100–122
	Not Elected	104.2	10.5	1.3	69–126
Otis IQ	Recommended	127.6	14.3	2.4	97–153
	Satisfactory	114.6	11.7	3.2	95–137
	Not Elected	101.8	13.6	1.6	74–130
Kuhlmann-Anderson IQ	Recommended	122.4	14.9	2.5	94–154
	Satisfactory	110.4	8.8	2.4	95–127
	Not Elected	103.5	9.5	1.1	82–127
Stanford Reading Grade Placement	Recommended	9.3	1.0	0.2	6.7–10+
	Satisfactory	8.4	1.0	0.3	6.3–9.9
	Not Elected	7.2	1.4	0.2	4.4–10+
Luria-Orleans Modern Languages Prognosis Score	Recommended	90.3	33.4	5.7	32–167
	Satisfactory	87.1	18.2	4.9	54–121
	Not Elected	50.0	24.3	2.9	10–109
Stanford Physiology Grade Placement	Recommended	8.2	1.5	0.3	5.2–10+
	Satisfactory	7.4	1.3	0.4	5.9–10+
	Not Elected	6.4	1.4	0.2	3.2–10+
Stenquist Mechanical Aptitude Score	Recommended	56.8	10.7	1.8	32–82
	Satisfactory	54.6	9.6	2.6	40–71
	Not Elected	58.8	11.9	1.4	28–78
Stanford Arithmetic Grade Placement	Recommended	8.3	1.3	0.2	6.0–10+
	Satisfactory	7.3	0.7	0.2	6.3–8.7
	Not Elected	6.7	1.2	0.1	4.3–10+
Orleans Algebra Prognosis Score	Recommended	78.8	33.4	5.7	33–163
	Satisfactory	57.9	26.6	7.2	35–126
	Not Elected	41.7	16.6	2.0	14–84

between the three intelligence tests used, and the reading grade placement is a close second as a basis of prediction.

Table I presents a detailed description of the test data segregated according to record in language. The "Recommended" group consists of the 34 students who elected a language and did work of superior quality, usually defined as A or B. The "Satisfactory" group consists of the 14 students who elected a language and did work of passing but not recommending quality, usually defined as C or D. The "Not Elected" is a heterogeneous group consisting of the 72 students who did not take a foreign language, either because it was unnecessary in their educational and vocational plans or because they showed no aptitude for it. Since there was only one failure, insufficient to establish any generalizations, that case was dropped from the tally.

There seems to be some systematic differentiation for all criteria except Stenquist Mechanical Aptitude score. The extent to which the differences are reliable is presented in Table II. The letter preceding the number indicates the superior group, and the probability column shows the chances in 100 that the obtained difference is a true difference.

In differentiating the Recommended from the Satisfactory group, the reliable tests in order of decreasing excellence are Terman, Kuhlmann-Anderson, Stanford Arithmetic, Otis and Stanford Reading. The Luria-Orleans Language Prognosis is the lowest on the list, perhaps because it claims to predict only for modern languages while data from Latin are included, thus decreasing the validity. The difference between the Recommended and Not Elected groups is greater, the reliable tests in order being Otis, Stanford Reading, Terman, Kuhlmann-Anderson, Luria-Orleans Language Prognosis, Orleans Algebra Prognosis, Stanford Arithmetic, and Chronological Age. Only the Stenquist fails to show a significant difference. The difference between the Satisfactory and Not Elected groups is reliably shown by the Luria-Orleans Language Prognosis, Stanford

TABLE II

Summary of Ratios

COMPARISON IN	RECOMMENDED AND SATISFACTORY				RECOMMENDED AND NOT ELECTED				SATISFACTORY AND NOT ELECTED			
	D	σD	Ratio	Prob.	D	σD	Ratio	Prob.	D	σD	Ratio	Prob.
Chronological Age...	S 1.7	1.6	S 1.1	S 86	N 5.2	1.2	N 4.3	N100	N 3.5	1.5	N 2.3	N 99
Terman IQ	R11.3	2.9	R 3.9	R100	R18.2	2.6	R 7.0	R100	S 6.9	2.2	S 3.1	S100
Otis IQ	R13.0	4.0	R 3.2	R100	R25.8	2.1	R12.3	R100	S12.8	3.6	S 3.6	S100
Kuhlmann - Anderson IQ	R12.0	3.4	R 3.5	R100	R18.9	2.8	R 6.8	R100	S 6.9	2.6	S 2.7	S100
Stanford Reading G.P.	R 0.9	0.3	R 3.0	R100	R 2.1	0.2	R10.5	R100	S 1.2	0.3	S 4.0	S100
Luria-Orleans Language Score	R 3.2	7.5	R 0.5	R 69	R40.3	6.4	R 6.3	R100	S37.1	5.7	S 6.5	S100
Stanford Physiology G.P.	R 0.8	0.4	R 2.0	R 98	R 1.8	0.3	R 6.0	R100	S 1.0	0.4	S 2.5	S100
Stenquist Mechanical Score	R 2.2	3.2	R 0.7	R 76	N 2.0	2.3	N 0.9	N 82	N 4.2	3.0	N 1.4	N 92
Stanford Arithmetic G.P.	R 1.0	0.3	R 3.3	R100	R 1.6	0.3	R 5.3	R100	S 0.6	0.2	S 3.0	S100
Orleans Algebra Score	R20.9	9.2	R 2.3	R 99	R37.1	6.0	R 6.2	R100	S16.2	7.5	S 2.2	S 99

Reading, Otis, Terman, Kuhlmann-Anderson, and Stanford Arithmetic.

In terms of the questions proposed, a general intelligence test was superior to the Luria-Orleans in two out of three discriminations and inferior in one. The same thing is true of the Stanford Reading Test. The best generalization seems that the three types of tests are equally valid, but the Luria-Orleans is somewhat more variable in result.

Though these data establish reliable differences between total groups, Table I also showed the great overlapping. Table III

TABLE III
Critical Scores

TEST OR TRAIT	PROBABLE NON-ELECTION	UNCERTAIN	PROBABLE RECOMMENDATION
Chronological Age	179–197	159–178	155–158
*Terman IQ	69–99	100–126	127–149
*Otis IQ	74–94	95–137	138–153
*Kuhlmann-Anderson IQ	82–93	94–127	128–154
*Stanford Reading G.P.	4.4–6.2	6.3–10+	?
*Luria-Orleans Language	10–31	32–121	122–167
Stanford Physiology G.P.	3.2–5.1	5.2–10+	?
Stenquist Mechanical Apt.	28–31	32–78	79–82
*Stanford Arithmetic G.P.	4.3–5.9	6.0–10+	?
Orleans Algebra Score	14–32	33–126	127–163

shows for the various tests the limits below which non-election seems advisable and above which superior accomplishment seems a certainty. The critical points are based on the ranges given in Table I. Tests marked by asterisks are those which, on the basis of Table II, seem of greatest validity in predicting achievement in foreign languages.

The major impression from such data is one of the great extent of overlapping between groups. It may be traced to the heterogeneity of the Non-Electing group, to the unreliability of the grades used as a criterion, or most probably to the mass

of factors other than innate ability which affect success such as personality, study habits, degree of motivation and the like. At any rate, any attempt to establish definite limits below which failure seems certain and above which success seems assured must be treated warily.

Because of the fact that most comparative data is in terms of coefficients of correlation rather than ratios it seemed desirable to convert the data into that form for comparative purposes. The results in terms of biserial r are presented in Table IV.

TABLE IV

Correlation with Course Records

RELATION IN	RECOMMENDED SATISFACTORY	RECOMMENDED NOT ELECTED	SATISFACTORY NOT ELECTED
Chronological Age	−.205 ± .125	−.473 ± .070	−.294 ± .103
Terman IQ	.532 ± .102	.775 ± .046	.372 ± .099
Otis IQ	.529 ± .102	.859 ± .036	.504 ± .091
Kuhlmann-Anderson IQ	.499 ± .104	.794 ± 0.44	.395 ± .098
Stanford Reading G.P.	.417 ± .112	.799 ± .043	.476 ± .093
Luria-Orleans Language	.065 ± .129	.735 ± .049	.760 ± .067
Stanford Physiology G.P.	.321 ± .119	.645 ± .057	.396 ± .098
Stenquist Mechanical	.127 ± .127	−.105 ± .084	−.201 ± .107
Stanford Arithmetic G.P.	.502 ± .104	.696 ± .053	.278 ± .104
Orleans Algebra Score	.366 ± .116	.776 ± .046	.461 ± .094

The interpretation as far as the relative merits of the various tests are concerned is the same as for Table II. In comparison with Kaulfers' summary, the median language prognosis is .73 or somewhat higher, the median of .46 for promise in algebra is very near his median for achievement in algebra, the median of .49 for achievement in reading is slightly higher, the median of .53 for intelligence tests is definitely higher, as is also the

median of .50 for achievement in arithmetic, while the median of – .29 for age is approximately the same. In general, it seems the relationships found are somewhat closer than those cited by Kaulfers, possibly a function of the narrower range of measuring instruments represented.

CONCLUSIONS

The following conclusions seem justified.

1. The Terman, Kuhlmann-Anderson and Otis Intelligence tests, the New Stanford Reading Test, the New Stanford Arithmetic Test, and the Luria-Orleans Modern Languages Prognosis test accurately differentiate degree of success in beginning foreign languages.

2. It is questionable whether the Orleans Languages Prognosis Test is superior to a general intelligence test in predictive value. The data are contradictory.

3. Reading achievement, though a reliable index of probable language record, is less valid than either general intelligence or the language prognosis test used.

4. Critical scores are of doubtful value because of the great extent of overlapping between groups.

5. The independence of the Stenquist Mechanical Aptitude Test is clearly demonstrated.

6. The total test picture modified by subjective estimates of personal factors is a better estimate of probable success than any single test.

A STATISTICAL ANALYSIS OF THE HUMM–WADSWORTH TEMPERAMENT SCALE[1]

BARBARA L. KRUGER
Columbia University

PURPOSE OF THE STUDY

THE purpose of this study is to discover through statistical analysis the interrelationships which exist among the components of temperament as measured by the Humm-Wadsworth Temperament Scale. Defining a component as "a group of traits" (4) the authors devised this test to determine an individual's temperament type on the basis of the relative strength of seven components which they have called normal, manic, depressed, autistic, paranoid, hysteroid, and epileptoid.

The field of personality testing is attracting the attention of laymen and professional psychologists, alike. There are many personality tests on the market, but few have been subjected to careful objective study and attempts at validation on the basis of recognized criteria.

We selected for study the Humm-Wadsworth Temperament Scale because it has been adopted by personnel directors in different parts of the country and is being used in large concerns as one of a battery of tests used in the selection of employees. So far the only literature on the scale is that written by the authors (1, 2, 4, 5, 6, 7). For this reason it seemed to offer a fertile field for research.

Since the scale purports to measure seven components of temperament which are distinguishable, each from the others, by characteristic attitudes and feelings, the first step in study-

[1] This study was made at the University of Southern California under the direction of Dr. Floyd Ruch and Dr. Donald Watt in the fall of 1937.

ing it seemed to be a statistical analysis of its validity in this respect. In other words, does the scale measure seven more or less distinct patterns or do two or more components overlap to such a degree that they are but two aspects of the same syndrome?

A DESCRIPTION OF THE HUMM-WADSWORTH TEST

Theory on which the test is based. As has been already stated, a component of temperament is regarded as a group of traits (4). According to Rosanoff, on whose theory of personality the scale is based, every individual possesses these components in some degree. Persons found in institutions for the insane and for criminals differ from normal or socially adjusted individuals in degree rather than in kind; that is, their manifestations of some predominant component of temperament have been uncontrolled and have functioned as liabilities rather than as assets.

The strength of the so-called "NORMAL" COMPONENT is especially important in the matter of adjustment to the demands of everyday life and in moments of stress and during periods of strain. The authors say "The 'Normal' Component is primarily a control mechanism, providing rational balance and temperamental equilibrium. It underlies the conservatism and conformity to socially acceptable conduct observed in the well adjusted subject."

The other components derive their names from psychopathology, and their traits are characteristic of the respective mental disorders.

The HYSTEROID COMPONENT, besides being associated with hysteria is identified with "criminalism." It is concerned essentially with self-preservation. It would appear that the authors were introducing an extraneous criterion when in describing this component they say, "An individual with an excess of this component possesses *a character defect with ethically inferior motivation,* manifested by malingering, stealing, lying, cheating, and similar anti-social behavior." From

the psychological standpoint, stealing, lying, and cheating are phenomena which are observed and dealt with as are temper tantrums and other devices employed by the individual attempting to control forces in his environment.

The CYLOID COMPONENT, associated with manic depressive psychoses and involutional melancholia, is considered in its two phases, the manic and the depressed. Both are "characterized by emotionality, fluctuations in activity and interference with voluntary attention." In scoring the answers to questions in the scale each phase is considered as a separate component.

The SCHIZOID COMPONENT, "characterized by heightened imagination," has two characteristic manifestations: autism and paranoia, each of which is regarded as a separate component of temperament.

The EPILEPTOID COMPONENT, which is associated with physical symptoms of epilepsy, latent if not manifest, is characterized by meticulous application to projects and by inspirations, often of a religious nature.

Development of the scale and the selection of questions. Accepting Rosanoff's theory of temperament, Dr. D. G. Humm, a consulting psychologist, and Mr. G. W. Wadsworth, Jr., a personnel manager, both of Los Angeles, California, developed this scale by making a statistical comparison of the answers made by normal and abnormal, or atypical, subjects to questions devised to detect the presence of the respective components.

The authors' detailed description (5) of how the questions were devised and selected is most interesting, not only in connection with this particular scale, but because of the light it throws on the matter of discovering significant questions for inclusion in any questionnaire of this sort. We will only state briefly the most salient points.

In order to determine which questions were discriminatory with regard to the presence or absence of a given component, Humm and Wadsworth tried them out on patients in mental

hospitals who were selected on the basis of their clinical records. "Those whose case studies furnished no evidence of marked possession of a given component were contrasted with subjects in whom the component was strongly manifested." An attempt was made to use only those suffering from unmixed psychoses. Inmates of the state prison were used as the group whose answers were significant with regard to the presence of the hysteroid component.

In order to determine which questions were diagnostic for the "Normal" component, individuals who had been committed to institutions were considered "abnormal" and their answers contrasted with those of persons who were employed and otherwise adjusted to everyday pursuits and whose case histories revealed no evidence of psychopathic tendencies (the normal group).

The function of the "no" count. Because so many subjects were used who were unwilling or unable to cooperate erratic responses were sometimes observed. These took the form in some instances of stereotyped or patterned responses, such as answering all questions, "yes," or all, "no," or of a given number, "yes," then a given number, "no." Among criminals, they took the form of leaving many significant questions unanswered. In order to ascertain whether the test had been validly taken the number of questions answered "no" were counted for each subject, and the authors considered as "valid subjects" only those whose "no" count fell between 138 and 196.[2]

Validity of the test. "The validity of the test was determined by a comparison of the test scores with facts in the case studies of the subjects, and was further determined by use of the test on subjects who had not been used in establishing the norms" (5).

[2] As the test is now administered a provision is made for giving such a subject a "second chance" by having him circle in red all those answers about which he is doubtful. These are then scored as "yes" if the count exceeds 196 and "no" if it is less than 138.

Reliability of the test. "Reliability was tested by the split-half method because it is known that subjects, especially of cycloid temperament are likely to be characterized by changes of mood; so that measurement at different times would introduce an unwarranted error" (6).

The reliability of the respective components ranges from .70 ± .03 for Schizoid Paranoid to .88 ± .01 for both "Cycloid Depressed" and "Schizoid Autistic."

Other considerations. We shall not go into the fine points of test administration, scoring, and the interpretation of scores because these are practical matters with which we are not concerned in this study. Our interest is in whether seven discrete components are discernible in a statistical analysis of test results.

DATA AND METHOD OF THIS STUDY

Sources of data. Hollerith cards,[3] in which component scores were already punched,[4] were secured for 437 males who had consulted the Adult Guidance Service, which is supervised by the Adult Education Section of the Los Angeles City Schools. This service is free and is available to the students attending the Adult Schools and to the adult public in general. Counseling assistance is given in "vocational, educational, business, personal, and health problems," according to a mimeographed announcement.

The group which furnishes our data, therefore, cannot be considered a cross section of the adult male population, but rather as representative of those many persons who for one reason or another have not found a satisfactory vocational adjustment. The majority are between 20 and 30 years old. We should expect their scores to differ in some important respects from those of a more heterogeneous group.

[3] For a description of the Hollerith machine and its equipment see: *Practical Applications of the Punched Card Method in Colleges and Universities*, New York, Columbia University Press, 1935.

[4] I wish to acknowledge my indebtedness to Mr. Robert C. Storement of the Guidance Service for supplying those cards and to the Director for permitting me to use them.

Treatment of data. Intercorrelations were worked out for each component with every other and with the "no" count.

RESULTS

Our findings are presented in Table 1.[5] They show a consistent negative relationship between the "Normal" component and each of the other six and a positive relation to the "no" count. Just as consistent is the negative relation of the latter to the scores for the components other than the "Normal."

The Epileptoid scores show only negligible correlations with all of the other component scores.

The Hysteroid scores show little relationship to these for other components, with the exception of the Paranoid.

The correlation between the Manic and Depressed components, the two phases of the Cycloid temperament, is very high (.87). In contrast to this is the low correlation of .30 between the Autistic and Paranoid components which are considered two phases of Schizophrenia.

The coefficient of .68 indicates a considerable degree of relationship between the Autistic and Depressed components.

A comparison of our findings with those presented by the authors of the test[6] and which are reproduced in Table II shows some striking resemblances.

The striking differences lie in our negative correlations of the normal with each of the other six components and in the finding of a considerable degree of relationship between the Paranoid and Depressed components and practically none between the Paranoid and Epileptoid components.

[5] We have not corrected the coefficients of correlation for attenuation because we wished them to be comparable to those published by the authors of the test. It should be kept in mind they would all be higher if so corrected.

[6] Besides these zero order correlations the authors of the test worked out partial correlations which, however, did not alter the relationships in most cases enough to suggest the advisability of our doing likewise.

TABLE I

Intercorrelations (Zero Order) of the Components of the Humm-Wadsworth Test Based on the Answers of 437 Males Who Consulted the Adult Counseling Service of Los Angeles

COMPONENT	NORMAL	HYSTEROID	MANIC	DEPRESSED	AUTISTIC	PARANOID	EPILEPTOID
"No" Count	.79	−.59	−.83	−.85	−.55	−.76	−.04
Epileptoid	−.22	−.07	.06	.15	.05	.05	
Paranoid	−.77	.56	.52	.54	.30		
Autistic	−.42	.02	.49	.68			
Depressed	−.75	.21	.87				
Manic	−.66	.27					
Hysteroid	−.53						

TABLE II

Correlations (Zero Order) between Responses of Subjects, Classified by Components, to All Valid Questions[7]

COMPONENT	NORMAL	HYSTEROID	MANIC	DEPRESSED	AUTISTIC	PARANOID
Epileptoid	.044	.177	.242	.056	.112	.406
Paranoid	.551	.531	.497	.168	.387	
Autistic	.438	−.194	.505	.635		
Depressed	.535	.032	.887			
Manic	.694	.187				
Hysteroid	.318					

FURTHER ANALYSIS

The negative correlations found by us between the "Normal" and the other six components of temperament as measured by this scale suggested the possibility that the group we were studying might be atypical with respect to the number having low profile scores on the "Normal" component.

In order to test this hypothesis the cards on which the scores were recorded were run through the "sorter" of the Hollerith

[7] *American Journal of Psychiatry*, Vol. 92, No. 1, July, 1935, p. 185.

TABLE III

An Analysis of the Group of 437 Males with Respect to the Distribution of Component Profile Scores

STRENGTH OF COMPONENT	NUMBER OF SUBJECTS SCORING IN EACH CATEGORY						
	Normal	Hysteroid	Manic	Depressed	Autistic	Paranoid	Epileptoid
3	31	11	138	99	29	63	0
2	64	28	85	46	34	82	2
1	56	48	98	108	131	39	16
0	128	154	61	79	70	100	105
-1	86	119	19	49	76	25	137
-2	32	56	25	28	43	70	131
-3	40	21	11	28	54	58	46
Total	437	437	437	437	437	437	437

machine, to determine for each component how many subjects in the group had a "plus 3" profile score,[8] how many a "plus 2," etc., through 0 which is considered "borderline" to "minus 3."

The results of this analysis are presented in Table III. They show very clearly that in our group only 151 individuals or 35 per cent might be classified as having a strong "Normal" component while 321 or 73 per cent have strong manic characteristics and 253 or 58 per cent show depressed symptoms. There were more subjects ranking as "strong" in Paranoid and Autistic components than in the Normal.

The differences between the authors' findings and ours with respect to the degree of relationship between the Paranoid component and the Epileptoid and Depressed components, respectively, may be a function of the difference between the two groups of subjects. The authors' group included a number of patients who had been committed to institutions because of epilepsy. Our group, on the other hand, included a large percentage of unemployed. It is quite possible that the constitutional inferiority of the epileptics gave rise to feelings which are characteristic of paranoia and that not being able to get a job made many members of our group depressed to an abnormal degree and also generated a sense of their being discriminated against.

CONCLUSIONS

1. When we take into consideration the composition of the group studied we may conclude that our findings are in agreement with the theory of the authors as to the importance of the relative size of the "Normal" component in determining adjustment. They say on page 6 of the *Manual of Directions*, "Very impressive among our findings was the matter of interrelationships. The same high quantitative showing (with the exception of the 'Normal' component) was made by a highly

[8] The profile values are derived from the raw score norms and are published in the Manual of Directions for scoring the test.

successful salesman as by a psychiatric subject suffering from a manic-depressive psychosis . . . its significance in terms of behavior differed by reason of the strength of the 'Normal' component which was high in the salesman and low in the psychiatric subject."

Table III reveals that many persons who exhibited strong tendencies toward behavior which in its extreme, uncontrolled form, is diagnosed as psychotic, did not reveal a correspondingly high degree of the "Normal." This may explain why the members of the group as a whole find themselves in situations in which they need the kinds of assistance offered by the Guidance Service.

A further indication that this may be the case is furnished by a comparison of Tables I and II, which reveals a consistently different relationship between the "Normal" component and each of the others for the two groups studied. For our group the coefficients are all negative, for the other group all positive.

2. The correlation between the "Normal" component and the "no" count being positive and those between the other components and the "no" count being negative, are in accord with the statement made by the authors of the test, that "the 'Normal' component is primarily a control mechanism . . . essentially a 'brake' or a 'balance wheel'" and "presents mainly characteristics associated with restraint."

3. Our intercorrelations for the components of temperament, with the exception of the "Normal," are not in accord with the theory that such identifiable syndromes exist:

a. The correlation between the Paranoid and Hysteroid components suggests overlapping of these two.

b. The Manic and Depressed components are so closely related as to suggest that as here measured they are one and the same. This high correlation may be due to the way that the weights for each question were derived. If the answers of subjects who were suffering from a manic-depressive psychosis were weighted for the Manic component without de-

termining whether the subjects were in the manic phase or the depressed we should expect to find that the final weights were a measure of the factors common to both rather than a means of distinguishing between the two patterns.

c. The correlations of the Autistic and Paranoid components with the others indicate that they do not exist as distinct factors in temperament or that the test, in its present stage of development, does not include sufficiently discriminating measures of these factors.

d. The above analysis of the differences between the findings for the two groups with respect to the size of the correlations between the Paranoid and the Epileptoid and between the Paranoid and the Depressed components, suggests that Paranoid reactions are the result of circumstances rather than being indicative of a distinct type of temperament.

4. There is no doubt that this test offers possibilities for further research, among them:

a. The determination of the reliability of the scale by the test-retest method. A scale which purports to measure temperament should stand up under the retest method. If it does not we must conclude that what it measures is something far more transitory, a "mood" for instance, rather than the fundamental personality pattern or organization which gives rise to such.

b. The determination of the validity of the scale by checking the replies of large numbers of persons, representative of a cross section of the community, against some objective criterion like a well devised objective rating scale which might include behavior-grams for the guidance of the rater.

BIBLIOGRAPHY

The Test Itself

(1) Humm, D. G., and G. W. Wadsworth. "A Diagnostic Inventory of Temperament. Preliminary Report." *Psychological Bulletin*, 1933, Vol. 30, p. 602 ff.

(2) ————————. "The Humm-Wadsworth Temperament Scale, Preliminary Report." *Personnel Journal*, 1934, Vol. 12, p. 314 ff.

(3) ————————. "The Humm-Wadsworth Temperament Scale." Published by Doncaster G. Humm, Los Angeles, 1934.

(4) ————————. "The Humm-Wadsworth Temperament Scale." *Manual of Directions.* Second Edition, 1935. Doncaster G. Humm, publisher.

(5) ————————. "The Humm-Wadsworth Temperament Scale." *American Journal of Psychiatry*, Vol. 92, No. 1, July, 1935, p. 163 ff.

(6) HUMM, D. G. "Outline for the Analysis of Temperament." Multigraphed Pamphlet.

(7) ROSANOFF, A. J. *Manual of Psychiatry*, Sixth Edition. New York: John Wiley and Sons, 1927.

STATISTICAL AIDS

(8) GARRETT, H. E. *Statistics in Psychology and Education.* Revised Edition. New York: Longmans, Green and Company, 1937.

(9) BAEHNE, G. W. *Practical Applications of the Punched Card Method in Colleges and Universities.* New York: Columbia University Press, 1935.

A SHORT SCORING METHOD FOR THE LINK P.Q. TEST

CHARLES C. GIBBONS
Ohio State University

THIS paper describes an economical method of scoring Link's P.Q. or Personality Quotient test. The publisher provides, with each test, a score sheet on which the individual's responses are entered and his score computed. The test manual states that hand scoring for Trait 1 alone takes about ten minutes; for all five traits, about twenty minutes. This scoring time has restricted the use of the Link test.

An attempt was made to find a time-saving method of scoring the Link test by the use of scoring strips and Veeder counters. The method by which these scoring strips were made is described below. By this method a test can be scored for Trait 1 in three minutes; for all five traits in eight minutes. It would be much less expensive for the publishers to print these scoring strips instead of the score sheets which they now provide. For the present, however, those who wish to use this method will need to make their own scoring strips. Two persons, working together, can make a complete set of scoring strips in about four hours.

HOW TO PREPARE SCORING STRIPS

To prepare the scoring strips, one needs a copy of the test and score sheet for boys, a copy of the test and score sheet for girls, and twenty-two strips of white cardboard, size $2\frac{1}{4} \times 11$ inches. It is more economical for two persons to work together in preparing the scoring strips. One person reads the score values from the score sheet which accompanies the test. The other person places a cardboard strip on a test and writes the

score values, as they are read, opposite the corresponding items.

Except for the first six questions, all items are scored either 0, 1, or 2. The score sheet provided with the test indicates the amount of credit to be given for a "Yes," "No," or "?" response. Making scoring strips consists of writing these score values on strips so that when the strip is laid on any test the score values will be in alignment with the corresponding test items. Figure 1 is a portion of a scoring strip for boys' Trait 1. The numbers in parentheses are the numbers of the items. The two columns on the key correspond to the two columns on each page of the test. No score values are entered for items 126, 145, 147, and 149, because the score sheet shows that these items are not scored for Trait 1. In this manner, strips are made for each of the four pages of the test.

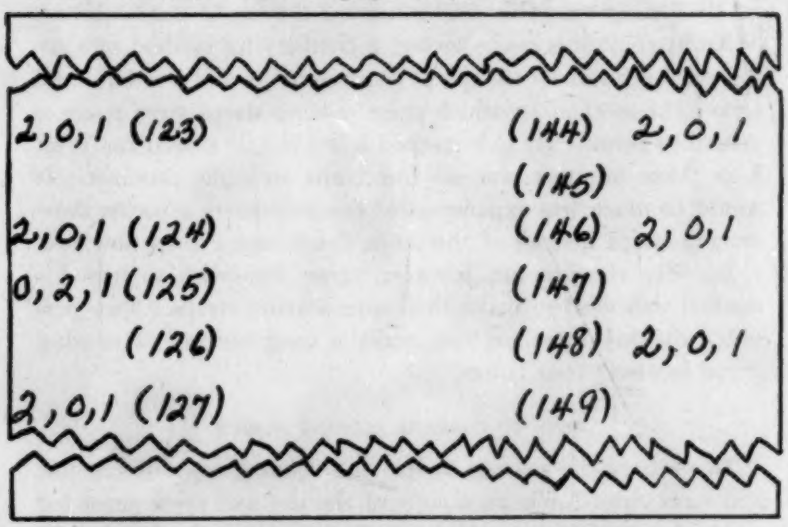


FIG. 1. Portion of scoring strip for boys' Trait 1.

To prepare scoring strips for the other four traits which the test attempts to measure, proceed in the same way except that the values for Traits 2 and 3 (Social Initiative and Self-Deter-

mination) are put on the same strips. Write Trait 2 values in red ink and Trait 3 values in blue ink. For the few items that are scored for both traits, write the values in red and encircle them with blue. In the same way, prepare a set of strips containing values for both Trait 4 and Trait 5 (Economic Self-Determination and Sex Adjustment). Write Trait 4 values in red ink and Trait 5 values in blue ink.

The scoring strips for girls' tests are prepared from a girls' test and answer sheet by the same method.

HOW TO USE THE STRIPS IN SCORING

For one who wishes to score a large number of Link tests, the following steps are suggested:

1. Separate the boys' tests from the girls' tests.

2. Secure two Veeder counters which have been mounted on a board and the four scoring strips for boys' Trait 1. For items 1–6, inclusive, compute the amount of credit by the rules stated on the score sheet (These rules may be abbreviated on a scoring strip). Count these values on the Veeder counter mounted on the right side of the board.

3. For the rest of the items, count on the right-hand counter the values on the scoring strip that correspond to the person's responses. The values 2, 0, 1 (see Fig. 1) mean 2 points credit for a "Yes" response, 0 points for a "No" response, and 1 point credit for "?." Similarly, 0, 2, 1 are the amounts of credit for "Yes," "No," and "?," respectively.

4. After all the papers have been scored for Trait 1, secure a set of strips for Traits 2 and 3. The method of counting values is the same as for Trait 1 except that values in red (Trait 2) are counted on the right-hand Veeder counter and values in blue (Trait 3) are counted on the left-hand Veeder counter.

5. Secure a set of strips for Traits 4 and 5. Proceed as before, counting values in red (Trait 4) on the right-hand counter and values in blue (Trait 5) on the left-hand counter.

Experience has shown that this scoring method: (1) Reduces scoring time for Trait 1 from 10 minutes to 3 minutes.

(2) Reduces scoring time for all five traits from 20 minutes to 8 minutes. (3) Eliminates the expense of having a score sheet for each person. (4) Decreases the likelihood of errors in scoring. (5) Reduces monotony and eye-strain on the part of the scorer.

NOTES

CRITICAL NOTE ON THE PRESSEY INTEREST–ATTITUDES TEST

ROBERT L. THORNDIKE
Columbia University

In 1933 Pressey and Pressey[1] reported a new personality test, which was an outgrowth of the earlier X–O test. This they called an Interest-Attitude test. It consists of four sub-tests of 90 items each. In the first test, the subject is instructed to check once everything that he thinks is wrong or that a person should be blamed for, and to check twice everything that he thinks is very wrong. The second test related to worries, the third to interests, and the fourth to traits admired in people.

The scoring of this test is in terms of a maturity score. Items are included in the test which show a progressive change in the frequency of checks as one goes from sixth-grade groups to junior and senior high school groups and on to college groups. Scores are to be translated into grade equivalents, and indirectly into age equivalents.

Certain items of the test show a greater frequency of checks in the higher age and grade group. In most cases, however, the frequency of checking decreases as one goes up the age scale. This is the case in 80 of the 90 items in the first two tests, 60 of the 90 in the third, and 40 of the 90 in the fourth, or in all 260 out of 360. This makes total score on the test, and especially score on the first two sub-tests, dangerously dependent upon what we may call the "checking threshold." That is, a person may get a low maturity score (corresponding to checking many immature and few mature items) merely because he tended to check freely both mature and immature items and the test gave more than twice as much opportunity to check the latter.

It is possible to demonstrate the influence of such a "checking threshold" by studying the responses on sub-tests 3 and 4, where the number of mature and immature items is more nearly equal. The test papers of a group of 25 gifted boys and a group of 24 gifted girls approximately 10–12 years of age were analyzed to bring out this factor. First, the total number of checks on mature items was correlated with the

[1] Sidney L. Pressey and Luella C. Pressey, "Development of the Interest-Attitude Tests," *Journal of Applied Psychology*, 1933, 17, 1–16.

number of checks on immature items. The correlation was .83 for boys and .90 for girls. In other words, there is a very marked positive relationship between the number of mature items checked and the number of immature items checked. If the test is measuring primarily maturity, we should expect the individual who marks many mature responses to mark few immature ones, and vice versa. This was not the case.

Then a correlation was obtained between score (immature responses minus mature responses) on one half of the items and the other half of the items. This was found to be .64 for boys and .75 for girls. In this case, the influence of the "checking threshold" is largely eliminated, as a result of the subtraction. We find that when this effect is eliminated there is still a fairly consistent differentiation on the basis of maturity.

Finally, a correlation was computed between total number of checks on one half of the mature and the immature items and total number of checks on the other half. The result was .87 for boys and .93 for girls. In our first pair of correlations the "checking threshold" and the maturity influence were working in opposition. In the correlations just reported, the influence of maturity is removed, and we get a pure indication of the consistency of the "checking threshold."

These results show clearly that for the group here studied an individual was characterized much more clearly and reliably by his "checking threshold" than he was by his maturity score. When the two forces were in opposition, as in our first correlations, the "checking threshold" was clearly predominant, and when the influences were separated, as they were in our second and third sets of correlations, the "checking threshold" provided the more reliable score.

For sub-tests 3 and 4, where the number of mature and immature items is about equal, the influence of the "checking threshold" is removed by the technique of scoring. In sub-tests 1 and 2, however, and in the total score for an individual, it may well be the predominant factor. Of course it is possible that mere willingness to check items may be a satisfactory index of immaturity. It seems, however, that we should view the score of an individual with scepticism when we find evidence that it may depend primarily on such a trivial and ill-defined tendency.

A NOTE ON THE CRITICAL NOTE

SIDNEY L. PRESSEY
Ohio State University

Dr. Robert Thorndike has kindly sent the writer a carbon of the note reporting the "checking threshold" as a major and largely irrelevant factor in scores of 48 children 10–12 years of age on the "Interest-Attitude" Tests. He much appreciates this opportunity to make early comment.

The possibility that mere tendency to check might be a dominant element in total response on such a test was so much a matter of concern to the writer when, nearly twenty years ago, he developed the X–O Tests that he put in every line of two of the tests a "joker" or word which would be marked only if the subject checked over-extensively and indiscriminately.[1] Marking of these jokers would signal the scorer when this factor appeared, and number of jokers marked might even serve as a measure of this tendency. But nobody seemed to make use of the jokers except the writer. And to his surprise, he found very few instances of such tendency to indiscriminate marking. In the X–O Form B the joker idea was not kept because of the above reasons and because emphasis was on the individual items and "net differential scores" where the factor should cancel itself out.

When work was started on the Interest-Attitude tests the matter was again given consideration. The original intention was to have in each test about an equal number of "mature" and "immature" items, so that the mere tendency to check would be cancelled out as a factor in total score. But extensive trial lists failed to yield sufficient wrongs and worries which were marked more often with increasing age because of increasing emotional maturity rather than increasing information or intelligence.[2] The empirical fact seemed to be that one important characteristic of the development of personality from ten to eighteen was a vigorous discarding of childhood ideas of wrong and childish fears.

[1] Pressey and Chambers, "First Revision of a Group Scale Designed for Investigating the Emotions," *Journal of Applied Psychology*, 1920, 4, 101.

[2] The point is briefly touched upon in the article describing the development of the tests (*Journal of Applied Psychology*, 17: 1–16) as is also the possibility of spurious effects due to the "tendency of some children to use a good many checks generally while some children tend to mark less."

To have dropped the wrongs and worries tests because of this lack of balance between mature and immature items would then have meant the omission of something very significant. A different test technique might have been used. But with this single-word question technique students were indicating three levels of response to 360 items (a total possible range of score of 720 points) in 30 minutes all on a small four-page folder, and scoring was easy (and fitted a plan for machine scoring which the author then had in mind). Convenience *is* worth while. A rich yield of score within a convenient time is needed for adequate measurement. Other techniques involved other possible irrelevant tendencies (most tests in the field involve some such). A change in technique at this juncture seemed likely to lose far more than it gained. The decision therefore was to build the tests on the data, and then see what results a trial form yielded.

The trial showed that from grades 5 to 12 medians changed 140 points—the score appeared sensitive to something. Further (contrary to what one would expect if mere tendency to check were an important factor), overlapping of grade groups was surprisingly slight for tests in this field; thus only 20 per cent of boys in the eighth grade scored below the median for the sixth. Comparisons of groups (as country high school versus sophisticated girls finishing school) showed results which would be expected. Correlations with estimated emotional maturity were reasonably good. Comparisons of item analyses with similar items in other studies of play and interest showed parallel curves.

It was therefore decided to issue the tests, with very complete data—norms for each sex and each test and for *each one of the 360 items* for a total of over 4000 cases. It was felt that the form should at least be valuable as a research questionnaire of exceptional range and richness of interpretive data (thus item comparisons between groups should not be grossly warped by such a general factor as the checking threshold). The scores *did* seem to have significance. And with both positive and negative items in each test (and in different proportions in different tests), and with such full data, an unusual variety of cross comparisons in research could be made, and possible faults readily detected. Presumably Dr. Thorndike's data at the very least leave the test this last merit!

The writer would argue that the merit is more important than might at first seem—and that Dr. Thorndike has touched upon a need far more important than he realizes. It is not after all an extraordinary omission in modern test technique that methods have not been developed by which the "internal consistency" of an individual's performance on a test may be readily determined![3] And is it not a merit of the Interest-Attitude

[3] Almost twenty years ago the writer argued that on the Binet irregularity might be used in something this way ("Irregularity on a Binet examination as a measure of its reliability," *Psychological Clinic*, 12: 236–40,

tests that they readily permit such handling? Since the first test contains only 10 negative items out of 90 and the last 50 out of 90, the difference between score on first and last tests should be evidence of the extent to which such factors as the "checking threshold" have operated. The writer is planning shortly to go over results now on file, in such fashion, and make further study of this total issue.

In the meantime he would contend (in summary) that the weight of evidence was against the conclusion that a "trivial and ill defined tendency to check" was often a devastatingly important factor in scores on these tests (though it may be so in such special groups of children, at the bottom of the age range for which the tests are planned*)—that the tests had, in fact, certain distinctive merits. He would argue further that Dr. Thorndike himself has all but demonstrated that a 'check upon checking' is readily obtained for each pupil. And he would contend finally that failure to utilize the interrelations between sub-scores as well as their sum, and to call for evidence of internal consistency as well as total in reporting an individual's score, are omissions in present-day testing methods which might well receive attention.

1919). But nobody (including the writer) made much use of the idea; and trial of the notion on group tests of intelligence yielded too inconclusive results (for reasons which need not be gone into here) to publish.

* However, in later correspondence Dr. Thorndike has kindly sent the scores of these children—and they score as more mature than their chronological age about as the writer would expect they should.

NOTE ON "I.Q. CHANGES IN FOSTER HOME CHILDREN" BY EMMETT L. SCHOTT

ROBERT L. THORNDIKE
Columbia University

In his report in this journal,[1] Schott states that the median gain of 5.8 points for 74 children retested after placement in foster homes is not statistically significant. He says "the ratio of the actual difference to the standard deviation of that difference is slightly less than .28."

A casual examination of the data suggested that this statement was probably incorrect, so the computations were checked using Schott's Table III. From this table it was found that the mean test-retest difference was +3.95 points of I.Q., the standard deviation of the distribution of differences was 7.3, and the standard error of the mean difference was 0.85. The difference was about 4.6 times its standard error. Obviously, it was statistically significant.

The conclusion of Schott's article should read: "After an average length of residence in a foster home of about one year, our group show a statistically reliable gain in I.Q."

[1] Emmett L. Schott, "I.Q. Changes in Foster Home Children," *Journal of Applied Psychology*, 1937, 21, 107–112.

NEWS AND NOTES

Within the past year a number of psychologists have conducted studies on problems pertaining to the radio under the stimulus and guidance of the Princeton Radio Research Project of the School of Public and International Affairs of Princeton University. These studies comprise a part of the research program in this field made possible by a grant from the Rockefeller Foundation and directed by Dr. Paul Felix Lazarsfeld in association with Dr. Hadley Cantril of Princeton, and Dr. Frank N. Stanton of the Columbia Broadcasting System.

Owing to the interest and importance of these studies the editors of the JOURNAL OF APPLIED PSYCHOLOGY are pleased to announce that its next issue, February, 1939, will be a special number devoted primarily to publication of these studies with Dr. Lazarsfeld as guest editor. The contents will include reports on the influence of the radio as a factor in different kinds of activities as well as research on fundamental index problems and problems of questionnaire handling. There will also be a review of the literature on listening habits and notes of current research in this general field being conducted at various institutions. The guest editor, Dr. Paul Felix Lazarsfeld, would appreciate communications from psychologists who have conducted, or are currently engaged in, any researches bearing on radio problems. He may be addressed at the Office of Radio Research, Field Headquarters, 22 East 17th Street, New York City.

The National Research Council announces the publication of a new quarterly entitled *Psychosomatic Medicine* with the sponsorship of the Committee on Problems of Neurotic Behavior, Division of Anthropology and Psychology. The inauguration of this publication has been made possible by a grant from the Josiah Macy, Jr., Foundation. The aim of *Psychosomatic Medicine* is to encourage and bring together studies which make a contribution to the understanding of the organism as a whole, in somatic and psychic aspects. These materials are now usually separated widely in manner and place of publication because of differences in concept, approach and methods. The board of editors include the following: Managing Editor, Flanders Dunbar; Editors, Franz Alexander, (Psychoanalysis); Dana W. Atchley (Internal Medicine); Stanley Cobb (Neurology); Hallowell Davis (Physiology); Flanders Dunbar (Psychiatry); Clark L. Hull (Psychology); Howard S. Liddell (Comparative

Physiology); and Grover F. Powers (Pediatrics). All correspondence relative to subscriptions or to the publication of articles should be addressed to Dr. Flanders Dunbar, Managing Editor, Academy of Medicine Building, Room 445, 2 East 103rd Street, New York City.

The University Tutorial Press, of London, England, now has ready a revised edition of "A Manual of Psychology" by G. F. Stout prepared in collaboration with C. A. Mace, of the University of London. Many chapters have been radically revised or largely rewritten, and a new chapter on "Practical Conative Activity" is introduced. An Appendix concerning Gestalt Psychology by Dr. R. H. Thouless has been included which should greatly increase the usefulness of the book.

The story of the production of copper is revealed in four new educational motion picture films made under the supervision of the Bureau of Mines, U. S. Department of the Interior, in cooperation with one of the larger mining companies. The titles of these films are "Copper Mining in Arizona," "Copper Leaching and Concentration," "Copper Smelting" and "Copper Refining," each showing the various steps in this interesting industry. Copies of these films, 16-mm. and 35-mm. size, for exhibition by those interested may be obtained from the Bureau of Mines Experiment Station, 4800 Forbes Street, Pittsburgh, Pa. No charge is made for the use of the films; the exhibitor pays only transportation charges. The width of film desired should be stated.

The Institute of Human Relations of Yale University, together with the American Association of Motor Vehicle Administrators and the Esso Safety Foundation from the Standard Oil Company of New Jersey, is announcing an automobile driver research project in an effort to study the human element in motor vehicle accidents and to establish uniform standards for the examination of motor vehicle operators. The project will be carried out under the immediate direction of Dr. Harry R. DeSilva, nationally recognized authority on highway safety and automobile drivers. The driver research project will be carried on primarily to obtain much needed material on which to base examinations of motor vehicle operators. It will also carry on studies into the relation between the ages of drivers a[illegible] accidents. Efforts will be made to determine the most efficient methods for reeducating careless drivers and to obtain new and scientifically accurate data on eye defects among motorists. It is hoped to unearth through this research many pertinent facts which will aid in the framing of new and more effective motor vehicle legislation and regulation. "Unfortunately," says Dr. DeSilva, "many state motor vehicle laws have been passed without the benefit of scientific facts ob-

tained from equally scientific research. This study will open up a whole new field. It will result in the setting up of minimum standards for the examination of applicants for licenses to drive motor vehicles and will go a long way toward the nation-wide standardization of the laws governing driver examinations in the various states."

The following new publications issued by the Office of Education, United States Department of Interior, may be obtained at nominal cost from the Superintendent of Documents, Government Printing Office, Washington, D. D.:

"Higher Education, 1930–36," "Statistics of City School Systems, 1935–36," "College Student Mortality," which tells how many students stop going to college, and why; "Education and the Civil Service in New York City," "Public Affairs Pamphlets—An Index to Inexpensive Pamphlets on Social, Economic, Political and International Affairs." How colleges and universities face or fail to face the problems of student accidents, illnesses, and fatigue is reported in a pamphlet, "Sanitation in Institutions of Higher Education." In "Nature and Use of the Cumulative Record" the author, Dr. David Segel, Specialist in Tests and Measurements in the Office of Education says, "Greatest aid to a pupil can be given only when his rate and trend of development in various intellectual, social and physical traits are known."

BOOK REVIEWS

HENRY C. LINK. *The Rediscovery of Man.* New York: Macmillan Company, 1938. 257 p. $1.75.

"Nero fiddled while Rome burned. Our popular leaders sing sweet songs while the character of a nation crumbles." These lines from the first chapter indicate not only the theme but also the scope of Dr. Link's book.

His thesis is that American personality and character are crumbling under the rising tide of paternalism, both in the home and in the government. The paternalism itself arises from three hundred years of concern with physical sciences and the consequent neglect of the capacities of man. Mankind's position and importance in the scheme of things has been impugned to such an extent that the individual is now prone to regard himself as a helpless clod tossed about and victimized by an environment over which he has no control. Governmental paternalism seems a rational solution of his difficulties to such an individual.

The Rediscovery of Man is aimed at the refutation of the view of the individual as irresponsible—a view which the author conceives to be mechanistic. Basing his approach on personality studies, Link takes into consideration habits essential to adequate, self-reliant personality development and the manner in which they may be acquired. The author has but little patience with the personality tinkering being done in fields outside psychology, pointing out that personality development and the modification of habits, desirable or undesirable, are strictly psychological problems. Link regards, and with justification, that non-psychological tinkering with personality is but another expression of the individual's wish to depend upon external agencies rather than take the responsibility for correcting his own traits. Communism, fascism, and our own benevolent paternalism with its Social Security fit into the same picture, he asserts.

The Rediscovery of Man is not a scientific book, as the term is properly used. It is pragmatic. Link is interested in getting action from his readers in the direction of assuming greater responsibility for their own welfare and personality characteristics, and in correcting in the future some of the present abnormalities. Probably few will agree with all the author says; most will be delighted that he has said it. Those who have a physical conception of organisms will be puzzled by Link's use of the term "mechanism" until they are well into the book where they will

discover that he is quite as "mechanistic" in his psychology as are most of us.

The book is written in a vivid and popular style, which is to say, the author knows his field sufficiently well to write about it in simple and exact words. No one will misunderstand what the book offers, but he will be provoked to intensive thinking in defense of some of his own beliefs.

This reviewer found the opening chapters a trifle slow and the central portion fascinating. The last part of the book, in which the author is applying his thesis to a wide range of current situations and beliefs and relating it to the Christian Ethic, seemed to lack Link's customary organization.

Psychologists have recently been crying for works that get into the home and the marketplace. *The Rediscovery of Man* gets there. Link has opened up a vital problem—so vital that one wonders if most psychologists are not also fiddling while Rome burns. The book may be read by all psychologists for both pleasure and profit. In the interest of humility and for the good of their scientific souls, the book is to be recommended particularly to those who, like the reviewer, hold a physical conception of psychology.

MATTHEW N. CHAPPELL,
342 Madison Avenue, New York City.

STRANG, RUTH. *Behavior and Guidance of Students in College and Secondary School.* New York: Harpers, 1937. Pp. 515. $4. *Counseling Techniques in College and Secondary School.* New York: Harpers, 1937. Pp. 159. $2.

These two volumes constitute the second and third of a series of five by the author on the general field of guidance in educational institutions. The emphasis is more on personal or personality guidance, although vocational aid is naturally bound up in the discussion. Volume II is chiefly concerned with adjustment. Strang opens with two lengthy chapters on the problems of the student, such as emotional maladjustment, disciplinary problems, vocational worries, and emancipation from the family. The next four chapters deal with measurement and interpretation of scores: intelligence, academic achievement, personality, and attitudes and interests. These contain comprehensive and critical summaries of these fields, worked out in such a way that a counselor is aided in the choice and utilization of the various tests and measures. Pitfalls are indicated throughout. Finally we have discussion of the social and economic background of students, the use of time by high school and college students, and financial problems of college students.

This volume must be considered a compendium more than anything else. Dr. Strang has worked in 1310 references, often as many as a dozen on a

single page, yet has skillfully blended in quotations and citations to make the manuscript smooth and readable, as well as informative. One's projected use of the book determines its valuation. To the well trained psychologist or counselor, much of the material will seem superfluous, as it duplicates topics taken up in general psychology, or courses in tests and measurements. But to the individual who may be asked to do some counseling on the side, in addition to his regular teaching duties, the book should be remarkably handy in giving much valuable data, critical outlooks, and a well oriented point of view.

Volume III aims to give more practical techniques of counseling. The principles suggested are equally applicable to college or high school levels. In the reviewer's opinion the outstanding feature of this shorter volume is the discussion of making the case study, handling the interview, making observations, and using rating scales. The various hidden angles, pitfalls, and traps into which the unwary counselor might fall are pointed out in very astute fashion. The necessity of keeping permanent records for cumulative purposes and to obviate probable loss of memory for details is well pointed out. The author has also pointed out the many variables in the background, as well as the present environment of the student which are necessary to consider in arriving at a comprehensive diagnosis and constructive solution of the problem, personal, educational, or vocational.

There is one limitation of this work, as well as of virtually every book on counseling, whether personal or vocational. This is the absence of a suggested constructive program. It is admittedly difficult to outline a sequence for someone else to follow; counseling must remain an individual matter, adapted to each individual, and hence it cannot be completely standardized. Yet, considering that Strang has given much elementary material which an experienced psychologist or trained counselor already knows, it would appear logical that the less experienced counselor would be very eager to find an outlined or skeleton organization of the conduct of a case from beginning to end, with approaches and possible modes of solution suggested. Possibly a few case studies, reproduced more or less verbatim from actual interviews, with objective data which are to be considered interspersed, would give constructive advice of value.

R. W. HUSBAND,
University of Wisconsin.

LELAND W. CRAFTS, THEODORE C. SCHNEIRLA, ELSA E. ROBINSON, RALPH W. GILCHRIST. *Recent Experiments in Psychology.* New York: McGraw-Hill Book Company, 1938. Pp. xiv + 417.

The diversity of what may legitimately be called psychological experimentation presents a serious problem in the choice of representative

studies. The authors of this volume have selected twenty-eight topics, devoting a chapter to each. These topics range from semi-popular studies of a "lightning-calculator" and of vocational analysis from photographs to the electro-physiology of the nervous system and the duplicity theory of vision.

In most chapters a brief introduction is followed by condensed reports of one or several studies. Purpose, method, results and conclusions are given in some detail for a majority of the experiments. Consideration of the statistical significance of results is avoided.

Recent Experiments in Psychology can perhaps be used with profit as supplementary reading for a class in experimental psychology. The condensation of journal material makes it too difficult for the student in an introductory course while the failure to consider the majority of evidence on any topic reduces its value for those capable of comprehending the content. In the opinion of the reviewer, the presentation of experimental method to the members of elementary courses in psychology demands emphasis upon the techniques of sound research rather than examples of investigation.

GEORGE K. BENNETT,
The Psychological Corporation.

LOUIS P. THORPE. *Psychological Foundations of Personality.* New York: McGraw-Hill Book Company, 1938.

It is a little difficult to decide exactly what audience was expected to receive this book. From the preface, one would infer that Dr. Thorpe was writing for undergraduate college students, particularly in the field of education, who were going to become school teachers. The preface states that, "This volume is offered as a text book in courses wherein the nature of that most basic of all psychological concepts, the human personality, is given primary consideration. It is also offered as a source of easily accessible information to that host of earnest educators and professional men . . . who are anxious to secure a background in this field."

After reading this preface, I expected to find a book which dealt in fairly simple terms with the general problems of personality, giving a considerable amount of practical advice as to personality development and training. On the contrary, I found the book to be written in a very technical fashion with only one chapter, about one-tenth of the book, which could be called practical in its approach.

Considered as a whole, the book suffers from excessive documentation. Dr. Thorpe reels up the impressive total of 1154 footnotes, which is probably a record of some kind. Unfortunately, many of these footnotes deal with relatively unimportant studies and are likely to confuse rather than

aid the undergraduate student. The author's manner of treating a given subject is to quote a considerable number of authors and investigations on each side of a controversial subject, finally giving his own conclusions. It was rather disappointing however, to notice that on some occasions his own conclusions could not be discovered and the controversial topic was left suspended in mid-air.

The introductory material is heavily loaded in the direction of a biological viewpoint; for example, some 30-odd pages are devoted to genes and the Mendelian studies of heredity—material which is probably rather technical for these students; but after having worked up this matter in such elaborate fashion, he is forced to say, "We must all guard against the apparently strong tendency to regard traits as unitary blocks in personality and as being potential or inevitable in gene inheritance." (Page 280.) If this is a tendency which must be guarded against, why should Dr. Thorpe go into such great detail about the mechanism of gene inheritance? Another instance of hereditary bias occurs in his discussion of maturation. Here he quotes those experiments which indicate that practice during the early years does not give a substantial advantage over maturation, but fails to quote any of the experimental data which show that training in certain types of activity may lead to substantial improvement which is not overcome by subsequent development.

I found myself most dissatisfied, however, with the treatment accorded to such very fundamental problems as feelings, emotions, and motives. His treatment of the psychoanalytic school is exceedingly short, and somewhat flippant. Freud, for example, gets a great deal less space in the book than MacDougall or Thorndike; although I feel reasonably certain that the science of psychoanalysis has contributed a great deal more to the understanding of personality than the educators who are frequently quoted. The treatment of Jung seems to me to be particularly unfair.

The author's point of view with regard to fundamental drives seems to be somewhat confused. At one time it seems that he has rejected the notion of motivational "energies" as having the same mystical nature as instincts. On page 275, however, we find this statement, "There is much as well to commend the idea of draining off energy, sexual or otherwise, in the physically active realms of work, play, and athletics."

Dr. Thorpe's position with regard to clinical psychology seems to be very critical. He makes the assertion, for example, that most clinical evidence is based too much on introspection and is lacking in experimental vertification. In his practical suggestions, however, we find that he relies not upon experimental, but upon clinical evidence.

If we consider the various books with which this one might be compared, we find that by dissipating itself in trying to cover too many different kinds of problems or points of view, it fails effectively to do any

of the various jobs which it might have done. From the point of view of objective and experimental psychology, it is somewhat inferior to Shaffer's *Psychology of Adjustment*, and from the point of view of psychological theory, it cannot compare with Allport's *Personality—A Psychological Interpretation*. It does not seem to me to have as many good practical suggestions, particularly for children, as Morgan's *Psychology of the Unadjusted School Child*. This evaluation may suggest that the book has no value at all, an impression which I did not intend to give. It does contain a wealth of material, but is presented in such fashion that graduate students will probably find it difficult enough and the average undergraduate student will probably find it impossible.

I should say that it would be useful as a handbook of psychological literature, particularly on the biological approach to many of these problems, because of its remarkable completeness of documentation.

A poor job of proof reading makes it a bit exasperating to read.

ROSS STAGNER,
University of Akron.

F. F. POWERS, T. R. MCCONNELL, W. C. TROW, B. V. MOORE, AND C. E. SKINNER. *Psychology in Everyday Living*. New York: D. C. Heath and Co., 1938. Pp. x+511. $2.75.

To the sophisticated reader the title of this book suggests that as a textbook it deviates from the conventional text, such as Woodworth's, Dashiell's, or Warren and Carmichael's, both in scope and in organization of subject matter. The book justifies its title. The authors have adhered to the express purpose of preparing a book "to meet the need of the college student for a well-integrated, modern text which will materially contribute to his ability to manage his own life. . . ."

A text that devotes more than one half of its space to personal and social adjustment and to "applications of psychology" naturally must do so at the expense of other subject matter. The most notable omissions are in the fields of the nervous system and of the sensory and perceptual processes. Approximately four pages are devoted to the nervous system in chapter 3, entitled Foundations of Behavior. In the same chapter slightly more than five pages are devoted to the receptor system. In chapter 14, Acquiring Knowledge and Skill, "sensing" merits nine pages and "perceiving" three. The latter type of organization may be defensible on the ground that sensing is essential to learning; yet it seems a bit illogical, if not misleading, to treat sense qualities in a chapter bearing this title. In addition to chapter 3 listed above, Part I, "The Nature of Psychology" (66 pages), contains two additional chapters, Psychology in Everyday Living, the introductory chapter, and Psychology as an Emerging Science.

Part II, "Understanding Ourselves" (164 pages), contains the following chapter headings: Personality, Psychology of Adjustment, Human Motivation, Our Emotional Life: Its Genesis and Control, Intelligence and Special Aptitudes, and Individual Differences. The last two, as well as the chapter entitled Psychology of Adjustment, are especially good; their treatment is unusually cogent and, in the case of the last two, ably supported by experimental data.

Part III, "Adjustment to College and Life" (75 pages), consists of the following three chapters: Social Behavior and Adjustment, Choosing a Vocation, and Psychology and Marriage. The writer feels that this section of the book is, in comparison with the others, rather weak. There is, of course, an obvious limitation in suitable material in this field. The authors responsible for this section have elsewhere in the text produced some of the ablest chapters.

Part IV, "Learning and Study" (83 pages), is made up of five chapters as follows: The General Nature of Learning, Acquiring Knowledge and Skill, The Acquisition of Attitudes, Learning to Study, and Problem-Solving and Creative Learning.

Part V, "Applications of Psychology" (91 pages), contains four chapters with the following titles: Psychology in Business and Industry, Psychology in Law, Politics, and Religion, Psychology in Medicine, and Psychology in Education.

The book contains half-tones of twenty eminent psychologists. Ten of these appear on coated stock on two pages opposite the title page, and ten appear in similar manner following page 4.

This book has the merit of being easily read. In this respect it excels most elementary texts. Neither the organization nor the style appears to have suffered from joint authorship, although the organization of the book is not one of its strong points. The reviewer feels that the authors have succeeded in a large measure in attaining their objectives. They have employed the available material to good advantage. That careful study of the book will be helpful to college students, most psychologists will probably agree. Perhaps many will agree that it will be more helpful than a study of the conventional text. Whether or not the content is adequate for the introductory course in psychology is another matter. Certainly it compares unfavorably with the standard text in the field with respect to what we have come to regard as the solid subject matter of psychology.

That the subject matter of the book under consideration has a useful place in the college curriculum the reviewer should agree. What its relationship to the introductory course in psychology should be, teachers of psychology will have to answer. This book and others like it may set the pattern for future courses, or they may find a useful place as collateral

reading in the introductory courses or may be incorporated into the subject matter of general orientation courses. We come back to the question of what psychology should be and who is to study it. Is it to the best interest of psychology to adapt it to popular demand? The student will probably gain little insight into psychological methodology or little appreciation of the scientific aspects of psychology from a study of a book of this kind. Moreover, one wonders about the advisability of devoting so little space to sensation and perception, the fields which have been developed more nearly to scientific status than have any of the others. If the psychologist feels that the introductory course should give a simple exposition of the fact and theory of psychology with emphasis upon its scientific aspects, he will be nettled by such omissions and will be unreceptive to such a chapter as Psychology and Marriage.

J. B. Stroud,
State University of Iowa.

NEW BOOKS AND PAMPHLETS RECEIVED

New books and pamphlets for review should be sent to James P. Porter, *Editor*, JOURNAL OF APPLIED PSYCHOLOGY, Ohio University, Athens, Ohio.

A Booklet of Norms. The Cooperative Test Service of the American Council on Education, 15 Amsterdam Avenue, New York, 1938. 85 pp.

Child Psychology. MARGARET WOOSTER CURTI. Longmans, Green & Company, New York, 1938. $3.20. 458 pp.

Child Psychology. An annotated bibliography. FLOYD HARDIN, EULALIA DOUGHERTY CHAPMAN AND LETHA BELLE HILL. Bibliographical Center for Research, Denver, Colorado, 1938.

Color-Music. THEODORE F. KARWOSKI AND HENRY S. ODBERT. Psychological Monographs, Vol. 50, No. 2. Psychological Review Company, Ohio State University, Columbus, Ohio, 1938. 60 pp.

The Conceptual Representation and the Measurement of Psychological Forces. KURT LEWIN. Psychological Theory, Vol. 1, No. 4. Duke University Press, Durham, N. C., 1938. $2.00. 247 pp.

The Human Problem in Schools. MARION MILNER. Methuen & Company, London, England, 1938. 8s 6d. 320 pp.

Logical Aspects of Educational Measurement. B. OTHANEL SMITH. Columbia University Press, New York, 1938. $2.50. 182 pp.

Recherches sur la Solidarite et L'Independance des Aptitudes Intellectuelles. HENRI LAUGIER AND D. WEINBERG. Imprimerie Chantenay, Paris, France, 1938. 79 pp.

Social Psychology. JAMES REINHARDT. J. B. Lippincott Company, New York, 1938. 467 pp.

Vocational Guidance in Catholic Secondary Schools. SISTER M. TERESA GERTRUDE MURRAY. Bureau of Publications, Teachers College, Columbia University, New York, 1938. $1.60. 163 pp.

What Church People Think about Social and Economic Issues. NORMAN L. TROTT AND ROSS W. SANDERSON. Association Press, New York, 1938. 79 pp.

INDEX OF SUBJECTS

Titles of books are followed by the name of the author of the work reviewed. All other references are to original articles.

Abilities, senescence, mental, 607.

Ability, college, significance, time, factor, tests, 192.

Academic success, measured, teachers' marks, value, Bernreuter Personality Inventory, direct, differential prediction, 576.

Achievement, elementary algebra, prediction, 493; ——, foreign languages, prediction, 632; ——, measures, personality, adjustment, scholastic, 597.

Acuity, ametropia, reading speed, relationship, measures, visual, 70.

Adjustment, scholastic achievement, measures, personality, 597; —— score card, comparison, standard, self-scoring form, Vineland, 87.

Adult women, technological, counseling, 455.

Advertisements, intelligence differences, reaction, 169.

Age, performance, repetitive manual task, relation, 295.

Ages, Alpha scores, regression formula, predict educational, 275.

Alertness measures, permanency, mental, 307.

Algebra, prediction, achievement, elementary, 493.

Alpha scores, regression formula, predict educational ages, 275.

Ametropia, reading, speed, relationship, measures, visual acuity, 70.

Analgesia contrasted, gas anaesthesia, novocain block, psychophysiological reactions, painful stimuli, hypnotic, 408.

Anesthesia, novocain block, psychophysiological reactions, painful stimuli, hypnotic analgesia contrasted, gas, 408.

Army Alpha, weighted scoring, primary factors, revision, 239.

Attitude test, critical note, Pressey interest-, 657.

Attitudes, department store, employee, 470; ——, Jewish, non-Jewish students, personality traits, 487; ——, marital similarity, socio-economic, 340; ——, training, construction, diagnostic scale, apprentice, 32.

Barometers, dependability, psychological brand, 1, 252.

Behavior and guidance of students in college and secondary schools (Strang), 667.

Bernreuter personality inventory, direct, differential prediction, academic success, teachers' marks, 576; —— personality inventory, simplified scoring method, 390.

Biological approach to the problems of abnormal behavior (Harrington), 445.

Block, psycho-physiological reac-

tions, painful stimuli, hypnotic analgesia contrasted, gas anaesthesia, novocain, 408.
Brand barometers, dependability, psychological, 1; —— barometers, validity, dependability, psychological, 252.

Capacities, performance, mental defectives, lace-making, 175.
Casualty insurance, relation, scores, Strong's interest analysis blanks, success, selling, 97.
Child guidance procedures (Institute for Juvenile Research), 321.
Children, Randall's Island performance, normal, 186.
Children, E. L. Schott, note, I.Q. changes, foster home, 662.
Clinics, field, driver, 59.
College, background, personal data, factors, prediction, scholastic success, 42; —— freshmen, sex differences, negro, 437; —— freshmen, study, vocational interests, teachers, 527; ——, significance, time, factor, tests, ability, 192.
Computation, intercorrelations, international scoring machine, 587.
Correctional school subjects, use, group intelligence tests, 504.
Counseling adult women, technological, 455.

Defectives, lace-making, analysis, capacities, performance, mental, 175; —— skilled, weaving, test pattern, mental, 385.
Definition of Psychology, The (Keller), 110.
Department store employee, attitudes, 470.
Detector, truth, lie, 50.
Diagnostic scale, apprentice attitudes, training, construction, 32.
Differences, intelligence, negro college freshmen, sex, 437; ——, reaction, advertisements, intelligence, 169.
Differential prediction, academic success, measured, teachers' marks, value, Bernreuter personality inventory, direct, 576.
Dominance-feeling, handwriting, non-dominant men, women, judging expressive movement, sex, 480; —— -feeling, phonograph records, voice, judging expressive movement, 620; ——, sales personality, social, 132.

Economic attitudes, marital similarity, socio-, 340; —— status, university students, aspects, socio-, 261.
Education of personality, the (Sufian), 110.
Educational ages, Alpha scores, regression formula, 275.
Elementary course, psychology, students want from, 282.
Estabrooks studiousness scale, analysis, Young-, 105.
Estimates, use of time, experiment, test reliability, 400.
Evidence, extra sensory perception, research undermined, 549.
Examination, weighted scoring, primary factors, revision, Army Alpha, 239.
Extra sensory perception, research undermined, evidence, 549.
Expressive movement, judgments, dominance-feeling, phonograph records, voice, 620; —— move-

ment, judgments, sex, dominance-feeling, handwriting, non-dominant men, women judging, 480.

Factor, tests, ability, college, significance, time, 192.
Factors, prediction, scholastic success, college, background, personal data, 42; ——, revision, Army Alpha examination, weighted scoring, primary, 239; ——, sales interview, psychological, 17.
Family position, personality, 347.
Feeling, handwriting, dominant, non-dominant men, women, judging expressive movement, sex, dominance-, 480; ——, phonograph records, voice, judging expressive movement, dominant-, 620.
Field, driver clinics, 59.
Foreign languages, prediction, achievement, 632.
Formula, predict educational ages, revised Alpha, regression, 275.
Foster home children, E. L. Schott, note, I.Q. changes, 662.
Freshmen, sex differences, intelligence, negro college, 437; ——, study, vocational interests, teachers college, 527.
Frölich's syndrome, superior intelligence, patients, 395.
Fundamentals of educational psychology (Douglas-Holland), 447.

Gas anaesthesia, novocain block, psycho-physiological reactions, painful stimuli, hypnotic analgesia contrasted, 408.
Graphic rating scale, house-mothers, -fathers, Mooseheart, 288.
Group intelligence tests, correctional school subjects, use, 504.

Handwriting, dominant, non-dominant men, women, judging expressive movement, sex, dominance-feeling, 480.
High school personnel work, personality schedule, 534.
Home children, E. L. Schott, note, I.Q. changes, foster, 662.
House-mothers, -fathers, Mooseheart graphic rating scale, 288.
Humm-Wadsworth temperament scale, statistical analysis, 641.
Hypnotic analgesia contrasted, gas anaesthesia, novocain block, psycho-physiological reactions, painful stimuli, 408.

Illusion, part-whole proportion, 421.
Industrial worker, the (Whitehead), 543.
Insurance, relation, scores, Strong's interest analysis blanks, success, selling casualty, 97.
Intelligence, negro college freshmen, sex differences, 437; ——, patients, Frölich's syndrome, superior, 395; —— tests, correctional school subjects, use, group, 504.
Intercorrelations, international scoring machine, computation, 587.
Interest analysis blanks, success, selling casualty insurance, relation, scores, Strong's, 97; —— -attitude test, critical note, Pressey, 657; —— blank, analysis, occupational scores, Strong vocational, 8; —— test, point-tally, method, scoring, Strong's vocational, 426.

Interests, teachers college freshmen, study, vocational, 527.
International scoring machine, rapid computation, tables, intercorrelation, 587.
Interview, psychological factors, sales, 17.
In the realm of mind (Myers), 224.
Introduction to child study (Strang), 543.
Inventory, direct, differential prediction, academic success, measured, teachers' marks, value, Bernreuter personality, 576; ——, psycho-somatic, 327; ——, simplified scoring method, Bernreuter personality, 390.
I.Q. changes, foster home children, E. L. Schott, note, 662.

Jewish, non-Jewish students, personality traits, attitudes, 487.
Judgments, sex, dominance-feeling, handwriting, dominant, non-dominant men, women, expressive movement, 480.

Lace-making, analysis, capacities, performance, mental defectives, skilled, 175.
Languages, prediction, achievement, foreign, 632.
Leading, readability, effects, 140.
Lie detector, truth, 50.
Link P.Q. test, short scoring method, 653.
Logic, theoretical and applied (Evans-Gamertsfelder), 114.

Machine, computation, intercorrelations, international scoring, 587; —— operating, prediction, success, power sewing, 350; —— operations, prediction, vocational proficiency, 518.
Manual task, relation, age, performance, repetitive, 295.
Marks, value, Bernreuter personality inventory, direct, differential prediction, academic success, measured, teachers', 576.
Measures, permanency, mental alertness, 307; ——, personality, adjustment, scholastic achievement, analysis, 597; ——, visual acuity, ametropia, reading speed, relationship, 70.
Men, one hundred superior, 367; ——, women, judging expressive movement, sex, dominance-feeling, handwriting, non-dominant, 480.
Mental alertness, measures, permanency, 307; —— defectives, lace-making, analysis, capacities, 175; —— defectives, skilled, weaving, test pattern, 385; —— testing, brief report, 211.
Mental conflicts and personality (Sherman), 545.
Movement, judgments, dominance-feeling, phonograph records, voice, judging expressive, 620.

Negro college freshmen, sex differences, intelligence, 437.
Non-dominant men, women, judging expressive movement, sex, dominance-feeling, handwriting, 480.
Non-Jewish students, personality traits, attitude, Jewish, 487.
Non-standardized tasks, analysis, capacities, mental defectives, lace-making, performance pattern, 175.

Normal children, Randall's Island performance series, 186.
Note, critical, 659.
Novocain block, psycho-physiological reaction, painful stimuli, hypnotic analgesia contrasted, gas anaesthesia, 408.
Nurses, selection, student, 119.

Objective analysis of musical performance: studies in the psychology of music (Seashore), 112.
Occupation, psychology, 320.
Occupational scores, Strong vocational interest blank, analysis, 8.
Operating, prediction, success, power sewing machine, 350.
Operations, prediction, vocational proficiency, machine, 518.

Painful stimuli, hypnotic analgesia contrasted, gas anaesthesia, novocain block, psycho-physiological reaction, 408.
Patients, Fröhlich's syndrome, superior intelligence, 395.
Pattern, mental defectives, skilled, weaving, test, 385.
Pediatrician in search of mental hygiene, a (Bronson), 221.
Perception, research undermined, evidence, extra sensory, 549.
Performance, repetitive manual task, relation, age, 295; —— -time, simple tasks, estimation, 79.
Personal data, factors, prediction, scholastic success, college, background, 42.
Personality: a psychological interpretation (Allport), 219.
Personality, adjustment, scholastic achievement, analysis, measures, 597; ——, family position, 347; —— inventory, direct, differential prediction, academic success, measured, teachers' marks, value, Bernreuter, 576; —— inventory, simplified scoring method, Bernreuter, 390; —— schedule, high school personnel work, use, short, 534; ——, social dominance, sales, 132; —— traits, racial, 271; —— traits, sales-person, 229.
Personnel form, restandardization, A–S reactions, 554; —— records, studies, reliability, 247; —— work, personality schedule, high school, 534.
Phonograph records, voice, judgments, dominance-feeling, 620.
Physiological reactions, painful stimuli, hypnotic analgesia contrasted, gas anaesthesia, novocain block, psycho-, 408.
Position, personality, family, 340.
Power sewing machine operating, prediction, success, 350.
P.Q. test, short scoring method, Link, 653.
Predict educational ages, Alpha scores, regression formula, 275.
Prediction, academic success, measured, teachers' marks, value, Bernreuter personality inventory, 576; ——, scholastic success, college, background, personal data, factors, 42.
Preference list, Michigan vocational, 558.
Pressey interest-attitude test, critical note, 657.
Primary factors, revision, Army Alpha examination, weighted scoring, 239.

Proficiency, machine operations, prediction, vocational, 518.
Proportion illusion, part-whole, 421.
Psychological brand barometers, dependability, validity, 252; —— brand barometers, reliability, 1.
Psychological foundations of personality (Thorpe), 669.
Psychology and life (Ruch), 323.
Psychology in everyday living (Powers, McConnell, Trow, Moore, Skinner), 671.
Psychology of advertising (Burtt), 111.
Psychology, students, want from elementary course, 282.

Randall's Island performance series, normal children, 186.
Rating scale, house-mothers, -fathers, Moosheart graphic, 288.
Reaction, advertisements, intelligence differences, 161.
Reactions, painful stimuli, hypnotic analgesia, gas anaesthesia, novocain block, psycho-physiological, 408; ——, personnel form, restandardization, A–S, 554.
Readability, effects, leading, 140.
Reading speed, relationship, measures, visual acuity, ametropia, 70.
Recent experiments in psychology (Crafts, Schneirla, Robinson, Gilbert), 668.
Records, studies, reliability, personnel, 247; ——, voice, judgment, dominance-feeling, phonograph, 620.
Rediscovery of man, the (Link), 666.
Reliability, estimates, use of time, experiment, test, 400; ——, personnel records, studies, 247; ——, psychological brand barometers, 1.
Reminiscence and rote learning (Ward), 223.
Repetitive manual task, relation, age, performance, 295.

Sales interview, psychological factors, 17; —— -person, personality traits, 229; —— personality, social dominance, 132.
Scale, analysis, Young-Estabrooks studiousness, 105; ——, apprentice attitudes, training, construction, diagnostic, 32; ——, housemothers, house-fathers, Mooseheart graphic rating, 288; ——, statistical analysis, Humm-Wadsworth temperament, 641.
Schedule, high school personnel work, short personality, 534.
Scholastic achievement, measures, personality, adjustment, 597; —— success, college, background, personal data, factors, prediction, 42.
School personnel work, personality schedule, high, 534; —— subjects, use, group intelligence tests, correctional, 504.
Score card, comparison, standard, self-scoring form, Vineland adjustment, 87.
Scores, regression formula, predict educational ages, Alpha, 275; ——, Strong's interest analysis blanks, success, selling casualty insurance, relation, 97; ——, Strong vocational interest blank, analysis, occupational, 8.
Scoring, Bernreuter personality inventory, simplified, 390; —— machine, rapid computation, in-

tercorrelations, 587; —— method, Link P.Q. test, short, 653, ——, primary factors, revision, Army Alpha examination, weighted, 239; ——, Strong's vocational interest test, point-tally, 426.
Self-scoring form, Vineland adjustment score card, standard, 87.
Selling casualty insurance, relation, scores, Strong's interest analysis blanks, success, 97.
Senescence, mental abilities, 607.
Sensory perception, research undermined, evidence, extra, 549.
Sewing machine operating, prediction, success, power, 350.
Sex, dominance-feeling, handwriting, dominant, non-dominant men, women, expressive movement, judging, 480.
Skilled weaving, test pattern, mental defectives, 385.
Social life and personality (Bogardus-Lewis), 448.
Socio-economic attitudes, marital similarity, 340; —— -economic status, university students, aspects, 261.
Somatic inventory, psycho-, 327.
Speed, relationship, measures, visual acuity, ametropia, reading, 70.
Spelling, difficulty, note, 211.
Standard, self-scoring form, Vineland adjustment score card, comparison, 87.
Status, university students, aspects, socio-economic, 261.
Stimuli, hypnotic analgesia contrasted, gas anaesthesia, novocain block, psycho-physiological reactions, 408.
Store, employee, attitudes, department, 470.
Strong's interest analysis blanks, success, selling casualty insurance, relation, scores, 97; —— vocational interest blank, analysis, occupational scores, 8; ——vocational interest test, point-tally, modified method, scoring, 426.
Student guidance techniques (Paterson, Schneidler, Williamson), 225.
Student nurses, selection, 119.
Students, aspects, socio-economic status, university, 261; ——, different vocations, comparison, women, 161; ——, personality traits, attitudes, Jewish, non-Jewish, 487; —— want from elementary course, psychology, 282.
Studiousness scale, analysis, Young-Estabrooks, 105.
Success, college, background, personal data, factors, prediction, scholastic, 42; ——, measured, teachers' marks, value, Bernreuter personality inventory, direct, differential prediction, academic, 576; ——, power sewing machine operating, prediction, 350; ——, selling casualty insurance, relation, scores, Strong's interest analysis blanks, 97.
Superior men, one hundred, 367.
Syndrome, superior intelligence, patients, Fröhlich's, 395.

Task, relation, age, performance, repetitive manual, 295.
Tasks, estimation, performance-time, simple, 79.
Teachers college freshmen, study, vocational interest, 527.
Teachers' marks, value, Bernreuter personality inventory, direct, dif-

ferential prediction, academic success, measured, 576.
Technological, counseling adult women, 455.
Temperament scale, statistical analysis, Humm-Wadsworth, 641.
Test, critical note, Pressey interest-attitude, 657; ——, point-tally, method, scoring Strong's vocational interest, 426; ——, reliability, estimates, use of time, experiment, 400; ——, short scoring method, Link P.Q., 653.
Tests, ability, college, significance, time, factor, 192; ——, correctional school subjects, use, group intelligence, 504.
Time, experiment, test, reliability estimates, use, 400; ——, factor, tests, ability, college, relative significance, 192.
Training, construction, diagnostic scale, apprentice attitude, 32.
Traits, attitudes, Jewish, non-Jewish students, personality, 487; ——, racial personality, 271; ——, sales-person, personality, 229.

University students, aspects, socio-economic status, 261.
Validity, dependability, psychological brand barometers, 252.
Vineland adjustment score card, comparison, standard, self-scoring form, 87.
Visual acuity, ametropia, reading speed, relationship, measures, 70.
Vocational interest blank, occupational, scores, Strong's, 8; —— interest, teachers college freshmen, study, 527; —— interest test, point-tally, modified method, scoring Strong's, 426; —— preference list, Michigan, 558; —— proficiency, machine operations, prediction, 518.
Vocations, comparison, women students, different, 161.
Voice, judgment, dominance-feeling, phonograph records, 620.

Wadsworth temperament scale, statistical analysis, Humm-, 641.
War dance: a study of the psychology of war (Howe), 449.
Weaving, test pattern, mental defectives, skilled, 385.
Weighted scoring, permanency, factors, revision, Army Alpha examination, 237.
What people want from business (Houser), 320.
Whole proportion illusion, part-, 421.
Women, judging expressive movement, sex, dominance-feeling, handwriting, non-dominant men, 480; —— students, different vocations, comparison, 161; ——, technological, counseling adult, 455.
Work, short personality schedule, high school personnel, 534.
Working with psychology (Ruch-Warren), 321.

Young-Estabrooks studiousness scale, analysis, 105.

INDEX OF AUTHORS

The names of authors of original contributions are printed in capitals and small capitals.

ABEL, THEODORA M., 175, 385.
Anderson, Amos C., 224, 447.
ANDERSON, ROSE G., 211, 455.
Avey, Albert E., 115.
BENNETT, GEORGE K., 390, 669.
BIJOU, S. W., 186.
BILLS, MARION A., 97.
BLANKENSHIP, ALBERT B., 112, 321, 325, 518, 544.
BROWN, RALPH, 408.
BRYAN, ALICE I., 161.
BURTT, HAROLD E., 17.
CANADY, HERMAN G., 437.
CHANNELL, RALPH, 59.
Chappel, Matthew N., 667.
COPELAND, HERMAN A., 247.
CORBIN, HORACE H., JR., 252.
DAVIDSON, CHARLES M., 307.
DE SILVA, HARRY R., 59.
DODGE, ARTHUR F., 132, 229.
Drake, Raleigh M., 114.
DROUGHT, NEAL, 597.
DUDYCHA, GEORGE J., 79.
DWYER, PAUL S., 8.
EAGLESON, ORAN W., 271.
Edgerton, Harold A., 226.
EISENBERG, PHILIP, 480, 620.
EMBREE, ROYAL B., JR., 504.
ENGLE, T. L., 534.
FISS, C. B., 119.
FOWLER, PHILIP, 367.
GEIGER, H. E., 32.
Gentry, J. R., 321.
GIBBONS, CHARLES, C., 653.
GREENE, EDWARD B., 558.
GREENLY, R. J., 32.
GUILFORD, J. P., 239.
HAMLIN, ROY, 175, 385.
HAYES, SAMUEL P., JR., 347.
Husband, Richard S., 668.
JENKINS, JOHN G., 1, 252.
KOHN, HAROLD A., 288.
KOLSTAD, ARTHUR, 470.
KOPAS, JOSEPH S., 426.
KRUGER, BARBARA, 641.
KUDER, G. FREDERIC, 587.
LAWTON, GEORGE, 607.
LEUBA, CLARENCE, 549.
Louttit, C. M., 222, 323, 543.
LUCKIESH, MATTHEW, 140.
MCFARLAND, ROSS A., 327.
MACVAUGH, GILBERT S., 275.
MITCHELL, GLENN E., 17.
MOSS, FRANK K., 140.
NEMZEK, CLAUDE, L., 576.
Oberlin, K. W., 111, 225.
OTIS, JAY L., 350.
PARTRIDGE, E. DE ALTON, 311.
PATERSON, DONALD G., 421.
PERL, RUTH E., 161.
PORTENIER, LILLIAN, 261.
PRESSEY, SIDNEY, L., 659.
REMMERS, H. H., 32.
REYMERT, MARTIN L., 288.
ROSLOW, SYDNEY, 448, 546, 554.
RUCKMICK, CHRISTIAN A., 50.
SAUM, ANNA LOIS, 192.
SCHILLER, NAOMI, 169.

Schott, Emmett L., 395.
Schultz, Richard, S., 554.
Scott, Carrie Mae, 42.
Seagoe, May V., 493, 632.
Seitz, Clifford P., 327.
Smith, Kinsley R., 295.
Smith, Mapheus, 400.
Stagner, Ross, 110, 221, 340, 449, 452, 671.
Stover, R. D., 119.
Stromberg, Elerov L., 70.
Stroud, J. B., 673.
Stuit, Dewey B., 527.
Sukov, May, 487.
Taylor, Howard R., 518.
Thorndike, R. L., 657, 662.
Tinker, Miles A., 421.
Tussing, Lyle, 282.
Vogel, Victor H., 408.
Wells, F. L., 367.
Williams, Ruth, 367.
Williamson, E. G., 105, 119, 487.
Wolf, S. Jean, 87.
Zalowitz, Eleanor, 620.

Directory of American Psychological Periodicals

(*Continued*)

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Two volumes every year. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5.00. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Provincetown, Mass.: The Journal Press. Subscription $7.50 per volume. One volume per year. Edited by Carl Murchison and an international cooperating board. Quarterly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1926.

Psychological Abstracts—Ohio State University, Columbus, Ohio: American Psychological Association. Subscription $7.00 a year ($7.25 foreign). 700 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of Genetic Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Journal of Social Psychology—Provincetown, Mass.: The Journal Press. Subscription $7.00. 500–600 pages annually. Edited by John Dewey, Carl Murchison, and an international cooperating board. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $3.00. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Educational Psychology—Baltimore, Md.: Warwick & York. Subscription $6.00. Edited by Jack W. Dunlap, Harold E. Jones, and Percival M. Symonds. Published monthly except June to August. Founded 1910.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Character and Personality—Durham, N. C.: Duke University Press. Subscription $3.00 per annum. Quarterly. Published simultaneously in Great Britain and the United States. Edited by Charles E. Spearman (London) and Karl Zener (Duke University). Founded 1932.

Journal of Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1000–1200 pages annually. Immediate publication for any accepted manuscript. Edited by Carl Murchison and an editorial board. Quarterly. The general field of psychology. Founded 1936.